

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: June 8, 2003, 15:31:12 ; Search time 7863 Seconds  
(without alignments)  
11636.681 Million cell updates/sec

Title: US-09-831-846-1  
Perfect score: 3144  
Sequence: 1 gcttggtccctctcgtga.....gcgggtttggattgcttatg 3144

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.ste.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sv.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3144	100.0	3144	9	AB033072	AB033072 Homo sapi
2	2927	93.1	3320	9	AB056799	AB056799 Macaca fa
C	1424.8	45.3	100426	9	AL591063	AL591063 Human DNA
3	1282	40.8	179820	2	AC011679	AC011679 Homo sapi
5	1166.6	37.1	159443	2	AC101792	AC101792 Mus muscu
C	1166.6	37.1	303563	2	AC124583	AC124583 Mus muscu
C	1166.6	37.1	308013	2	AC125047	AC125047 Mus muscu
8	1151.4	36.6	197893	2	AC127880	AC127880 Rattus no
C	827.4	26.3	197893	2	AC127880	AC127880 Rattus no
10	634.6	20.2	1887	6	AX179258	AX179258 Sequence
11	634.6	20.2	2549	9	BC003378	BC003378 Homo sapi
12	633	20.1	1887	6	AX266997	AX266997 Sequence
13	633	20.1	2558	6	AX266995	AX266995 Sequence
14	621	19.8	2458	9	BC015581	BC015581 Homo sapi
15	608	19.3	1653	6	AX358336	AX358336 Sequence
16	587.8	18.7	2612	10	BC023156	BC023156 Mus muscu
17	563.8	17.9	3174	9	AB040917	AB040917 Homo sapi
18	554.6	17.6	1653	6	AX358338	AX358338 Sequence
19	549.2	17.5	3047	9	AK055365	AK055365 Homo sapi
20	544.4	17.3	3701	9	AK096627	AK096627 Homo sapi
C	516.4	16.4	180153	2	AC018798	AC018798 Homo sapi
22	511	16.3	139072	9	AP000485	AP000485 Homo sapi
23	511	16.3	182769	9	AP003176	AP003176 Homo sapi
24	477.2	15.2	92797	9	AF038458	AF038458 Homo sapi
C	472	15.0	138538	2	AC011445	AC011445 Homo sapi
C	472	15.0	172090	2	AC018477	AC018477 Homo sapi
27	465	14.8	71605	2	AC119332	AC119332 Rattus no
28	462.6	14.7	1889	9	BC027475	BC027475 Homo sapi
C	447.8	14.2	173649	2	AC117169	AC117169 Rattus no
C	447.6	14.2	118112	9	HS462C17	AL033380 Human DNA
31	446.4	14.2	1683	10	BC004018	BC004018 Mus muscu
32	438	13.9	168508	2	AC087141	AC087141 Mus muscu
33	438	13.9	198578	2	AC087138	AC087138 Mus muscu
C	432.6	13.8	279589	2	AC073768	AC073768 Mus muscu
C	403.6	12.8	158885	9	CNS01DWT	AL138498 Human chr
36	401.4	12.8	4301	9	HS803436	AL832129 Homo sapi
37	346.6	11.0	1832	9	BC025310	BC025310 Homo sapi
C	308.6	9.8	138034	2	AC096476	AC096476 Rattus no
C	305.6	9.7	205026	2	AC106165	AC106165 Rattus no
40	263.6	8.4	2271	9	BC014678	BC014678 Homo sapi
C	252.2	8.0	100660	2	AC095875	AC095875 Rattus no
C	250.4	8.0	63949	2	AC110498	AC110498 Mus muscu
43	213.2	6.8	1307	9	BC014040	BC014040 Homo sapi
44	208	6.6	1286	9	BC000207	BC000207 Homo sapi
45	204.2	6.5	432	6	AX358333	AX358333 Sequence

ALIGNMENTS

RESULT 1  
AB033072  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE

AB033072  
Homo sapiens mRNA for KIAA1246 protein, partial cds.  
AB033072.1 GI:6330832

Homo sapiens brain cDNA to mRNA, clone lib:pBluescriptII SK plus  
Clone:h00149a.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS  
Nagase,T., Ishikawa,K., Kikuno,R., Hirotsawa,M., Nomura,N. and  
Ohara,O.

TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	TITLE	JOURNAL
1. <u>Effect of the</u>						
2. <u>of the</u>						
3. <u>of the</u>						
4. <u>of the</u>						
5. <u>of the</u>						
6. <u>of the</u>						
7. <u>of the</u>						
8. <u>of the</u>						
9. <u>of the</u>						
10. <u>of the</u>						
11. <u>of the</u>						
12. <u>of the</u>						
13. <u>of the</u>						
14. <u>of the</u>						
15. <u>of the</u>						
16. <u>of the</u>						
17. <u>of the</u>						
18. <u>of the</u>						
19. <u>of the</u>						
20. <u>of the</u>						
21. <u>of the</u>						
22. <u>of the</u>						
23. <u>of the</u>						
24. <u>of the</u>						
25. <u>of the</u>						
26. <u>of the</u>						
27. <u>of the</u>						
28. <u>of the</u>						
29. <u>of the</u>						
30. <u>of the</u>						
31. <u>of the</u>						
32. <u>of the</u>						
33. <u>of the</u>						
34. <u>of the</u>						
35. <u>of the</u>						
36. <u>of the</u>						
37. <u>of the</u>						
38. <u>of the</u>						
39. <u>of the</u>						
40. <u>of the</u>						
41. <u>of the</u>						
42. <u>of the</u>						
43. <u>of the</u>						
44. <u>of the</u>						
45. <u>of the</u>						
46. <u>of the</u>						
47. <u>of the</u>						
48. <u>of the</u>						
49. <u>of the</u>						
50. <u>of the</u>						
51. <u>of the</u>						
52. <u>of the</u>						
53. <u>of the</u>						
54. <u>of the</u>						
55. <u>of the</u>						
56. <u>of the</u>						
57. <u>of the</u>						
58. <u>of the</u>						
59. <u>of the</u>						
60. <u>of the</u>						
61. <u>of the</u>						
62. <u>of the</u>						
63. <u>of the</u>						
64. <u>of the</u>						
65. <u>of the</u>						
66. <u>of the</u>						
67. <u>of the</u>						
68. <u>of the</u>						
69. <u>of the</u>						
70. <u>of the</u>						
71. <u>of the</u>						
72. <u>of the</u>						
73. <u>of the</u>						
74. <u>of the</u>						
75. <u>of the</u>						
76. <u>of the</u>						

gene  
CDS

gene  
CDS

BASE COUNT  
ORIGIN

Qy	1441	GATGACCGCCTGGTAGGGAAC	TCTCTAAGGACCGCTGTCTATGACAAATGCAACCTTGGAC	1500
Db	1441	GATGACCGCCTGGTAGGGAAC	TCTCTAAGGACCGCTGTCTATGACAAATGCAACCTTGGAC	1500
Qy	1501	ATCTTTATCAACCATCTCAGGACAGTGTCCCTTCACTGCTGCAATGCTGCGCAATGCTGCC	1560	
Db	1501	ATCTTTATCAACCATCTCAGGACAGTGTGTGCTTCACTTGCAATGCTGCGCAATGCTGCC	1560	
Qy	1561	GGAGAGCCACGGCCATGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGC	1620	
Db	1561	GGAGAGCCACGGCCATGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGC	1620	
Qy	1621	ACNAGCCGACTGCACCCCCAAGTCCCGCTCTCAGACATCACTGGCTCCAGCAAGACC	1680	
Db	1621	ACCAGCCGCACTGCACCCCCAAGTCCCGCTCTCAGACATCACTGGCTCCAGCAAGACC	1680	
Qy	1681	AGCCGGGAGGTGGAGGACGTGGGGGCGAGAGCCTCCCAAAGCCCGGAAACGGCT	1740	
Db	1681	AGCCGGGAGGTGGAGGACGTGGGGGCGAGAGCCTCCCAAAGCCCGGAAACGGCT	1740	
Qy	1741	GTGCTTGCTGTAAGTGAACAACACTCGGCCCTGGTCAAGTGGTGTGTGACGAAGTCA	1800	
Db	1741	GTGCTTGCTGTAAGTGAACAACACTCGGCCCTGGTCAAGTGGTGTGTGACGAAGTCA	1800	
Qy	1801	GCACCCGGGTGAAGATGTACCAGCTCCAGTACAACTGCTCTGACGATGAGGTACTGATT	1860	
Db	1801	GCACCCGGGTGAAGATGTACCAGCTCCAGTACAACTGCTCTGACGATGAGGTACTGATT	1860	
Qy	1861	TACAGGATGATCCAGGCTCCAAACAAGCCCTTGGTGTCAACAACCTGGTGTGACGGGACT	1920	
Db	1861	TACAGGATGATCCAGGCTCCAAACAAGCCCTTGGTGTCAACAACCTGGTGTGACGGGACT	1920	
Qy	1921	GGCTACGACTTGTGTGCTGGCCATGTGGATGACACAGCCACGACACTCACGGCCACC	1980	
Db	1921	GGCTACGACTTGTGTGCTGGCCATGTGGATGACACAGCCACGACACTCACGGCCACC	1980	
Qy	1981	AACATCGTGGCTCGGCCAGTCTTTCACCAAGCTGACTACCGGAGTGCAGTCCATG	2040	
Db	1981	AACATCGTGGCTCGGCCAGTCTTTCACCAAGCTGACTACCGGAGTGCAGTCCATG	2040	
Qy	2041	CACAGCCAGATTCTTGGCGGCAACCATGATCCTGGTCAATCGGGGGCATCATCTGTGGCCAGC	2100	
Db	2041	CACAGCCAGATTCTTGGCGGCAACCATGATCCTGGTCAATCGGGGGCATCATCTGTGGCCAGC	2100	
Qy	2101	CTGCTGTCTTCACTGCTATCTCATGGTGGCTTACAAGTCTTGCAACGAGGCCCC	2160	
Db	2101	CTGCTGTCTTCACTGCTATCTCATGGTGGCTTACAAGTCTTGCAACGAGGCCCC	2160	
Qy	2161	AGCAAGATGACGCGCGGTGAGCAATGTGACTCGCAGACCAACGGCGCCAGCCACCG	2220	
Db	2161	AGCAAGATGACGCGCGGTGAGCAATGTGACTCGCAGACCAACGGCGCCAGCCACCG	2220	
Qy	2221	CCTCCAAGCAGCCACAGCCGGGGCCCCCGCAGGGCCCCCGCGAGGTGGTGGCGC	2280	
Db	2221	CCTCCAAGCAGCCACAGCCGGGGCCCCCGCAGGGCCCCCGCGAGGTGGTGGCGC	2280	
Qy	2281	AACGAGTCTTGACTTCAACCGCAGCTGGCGCGGCCAGTGACTCTCTTCTCTCAGC	2340	
Db	2281	AACGAGTCTTGACTTCAACCGCAGCTGGCGCGGCCAGTGACTCTCTTCTCTCAGC	2340	
Qy	2341	TCCCTGGGAGTGGGAGGTGGGGGCTGGGACGGGCCCTTGGAGGATCCACACCTCC	2400	
Db	2341	TCCCTGGGAGTGGGAGGTGGGGGCTGGGACGGGCCCTTGGAGGATCCACACCTCC	2400	
Qy	2401	GCCCGGGCCCCAAGCCAGCCTTGAACCGCTGATGGGGCTTGGCTCTCTCTGGACCTC	2460	
Db	2401	GCCCGGGCCCCAAGCCAGCCTTGAACCGCTGATGGGGCTTGGCTCTCTCTGGACCTC	2460	
Qy	2461	AAGAGTCAGAGAAAGGAGAGTCTCTGACTCAGGACTCAGCCGGGAGAGGGGTGGG	2520	
Db	2461	AAGAGTCAGAGAAAGGAGAGTCTCTGACTCAGGACTCAGCCGGGAGAGGGGTGGG	2520	

QY	2521	ACGTGCGCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGG	2580
DB	2521	ACGTGCGCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGG	2580
QY	2581	GCCAGGAGCCTGCTCCCTTGCCTTGCAGGGCAAGGCCAAAACGACGCACTCCTTCGCAC	2640
DB	2581	GCCAGGAGCCTGCTCCCTTGCCTTGCAGGGCAAGGCCAAAACGACGCACTCCTTCGCAC	2640
QY	2641	ATGGGGACATTTGCTGCTGCGCGCGGGGAGGGTCTGTCCGGCGGCTACAGTCCTCCT	2700
DB	2641	ATGGGGACATTTGCTGCTGCGCGCGGGGAGGGTCTGTCCGGCGGCTACAGTCCTCCT	2700
QY	2701	CGGAAGGTCGCAACATCTGGAACGAAGCGAGCCTCTCTGTCAACGCGCATGCTCTTGCCC	2760
DB	2701	CGGAAGGTCGCAACATCTGGAACGAAGCGAGCCTCTCTGTCAACGCGCATGCTCTTGCCC	2760
QY	2761	TTTCAGGAGAGTGACCTGGTGGGGCCCCGGGGAGCTTTTGGACGCTCCGAATGGGTGATG	2820
DB	2761	TTTCAGGAGAGTGACCTGGTGGGGCCCCGGGGAGCTTTTGGACGCTCCGAATGGGTGATG	2820
QY	2821	GAGAGCAGGTCFAGGTGGGGGTGGGCAATGCTCCCTTTCTGTGCGCAGGGTGGGAGAAG	2880
DB	2821	GAGAGCAGGTCFAGGTGGGGGTGGGCAATGCTCCCTTTCTGTGCGCAGGGTGGGAGAAG	2880
QY	2881	GGGAAAGAACTCTACTGCGAAGTGTTTGTGGAGTTTCCATGGTCATGTTTACATCCAGGG	2940
DB	2881	GGGAAAGAACTCTACTGCGAAGTGTTTGTGGAGTTTCCATGGTCATGTTTACATCCAGGG	2940
QY	2941	ACAGTTTTCGTCCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCA	3000
DB	2941	ACAGTTTTCGTCCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCA	3000
QY	3001	CCTCCCAACACCCCGCGGGGTGTCTCAGGGAATGTGACCTCGCTCAAAATGCCGACT	3060
DB	3001	CCTCCCAACACCCCGCGGGGTGTCTCAGGGAATGTGACCTCGCTCAAAATGCCGACT	3060
QY	3061	GAGCCCTGAGTGTGTGGAAAGGCGAGACTCCGCCCTTTCTAATCACAAATGTAGCCTCAAA	3120
DB	3061	GAGCCCTGAGTGTGTGGAAAGGCGAGACTCCGCCCTTTCTAATCACAAATGTAGCCTCAAA	3120
QY	3121	GCAAGCGCCTTTGGATTGCTTATG	3144
DB	3121	GCAAGCGCCTTTGGATTGCTTATG	3144

RESULT 2	AB056799	3320 bp	linear	PRI 14-MAR-2001
LOCUS	AB056799			
DEFINITION	Macaca fascicularis brain cDNA clone:Qf1A-11865, full insert sequence.			
ACCESSION	AB056799			
VERSION	AB056799.1	GI:13365898		
KEYWORDS	fis (full insert sequence); oligo capping.			
SOURCE	Macaca fascicularis adult male frontal lobe left cDNA to mRNA, clone lib:macaque brain cDNA library Qf1A clone:Qf1A-11865.			
ORGANISM	Macaca fascicularis			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.			
REFERENCE	1 (sites)			
AUTHORS	Osada, N., Hida, M., Kusunda, J., Tanuma, R., Iseki, K., Hirai, M., Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.			
TITLE	Isolation of full-length cDNA clones from macaque brain cDNA libraries			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 3320)			
AUTHORS	Hashimoto, K., Osada, N., Hida, M., Kusunda, J. and Sugano, S.			
TITLE	Direct Submission			
JOURNAL	Submitted (09-MAR-2001) Katsuyuki Hashimoto, National Institute of			

Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan  
(E-mail: khashi@nih.go.jp, URL: <http://www.nih.go.jp/yoken/genebank/>,  
Tel: 81-3-5285-1111 (ex. 2120) Fax: 81-3-5285-1181)

COMMENT	Lab host: TOP10	QY	299	CATCTGCTGCAATT-----CGCCGAGCCCGCCGAGGGTTAGCTCG	340
	Vector: pME18S-PL3 (Acc.No. AB009864)	Db	379	CATCTGCTGCTGCTTCGACAGCCGCGGAGCCGCGGAGGGTTAGCTCG	438
	R. Site1: DraIII (CACTGTGTG)	QY	341	AGTCTGTCTCGGGGGGAGGATGCTGCGAGCCGCGGAGCCCGGCGCCCGCGGA	400
	R. Site2: DraIII (CACCATGTG)	Db	439	AGTCTGTCTCGGGGGGAGGATGCTGCGAGCCGCGGAGCCCGGCGCCCGCGGA	498
	Description: 1st strand cDNA was primed with an oligo (dT) primer using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-PL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by Sugano et al. (University of Tokyo, Institute of Medical Science). Custom primer used for sequencing	QY	401	GCGGCTCTGGTGCACCCAGCCGGGGGTAGTCTGCTCGCCCGCCAGGC-----GCTG	453
	( 5' end primer [CTTGTGCTCTAAAGCTGCG];	Db	499	GCGGCTCTGGTGCCTCCAGCCGGGGGTAGTCTGCTCGCCCGCCAGGCACCGTGGCCTG	558
	3' end primer [CGACCTGCAGCTCGACACA] ).	QY	454	AGTGACAGACCAATGGAGACCTCTGCTGGTGGCTGCTAGCGTTGGCATGGGTTTGGC	513
	Location/Qualifiers	Db	559	AGTGACAGACCAATGGAGACCTCTGCTGGTGGCTGCTAGCGTTGGCATGGGTTTGGC	618
	1. 3320	QY	514	GTGCTGACGCTGCCCCCAAGTACTGTCTGCCAGAACTCTGTGAGTCACTGGGAC	573
	/organism="Macaca fascicularis"	Db	619	GTGCTGACGCTGCCCCCAAGTACTGTCTGCCAGAACTCTGTGAGTCACTGGGAC	678
	/db_xref="taxon:9541"	QY	574	CTGTGCCCCCTCAAAGGGCTGCTCTTTGTATCCCCCTCATATTGACCGCGGACAGTGGAG	633
	/clone="Qf1A-11865"	Db	679	CTGTGCCCCCTCAAAGGGCTGCTCTTTGTATCCCCCTCATATTGACCGCGGACAGTGGAG	738
	/sex="male"	QY	634	CTGCGCTGCGCGGCAACTTCATCATCAATCAGCCGCGGAGACTTTGCCAATGACG	693
	/tissue type="frontal lobe left"	Db	739	CTGCGCTGCGCGGCAACTTCATCATCAATCAGCCGCGGAGACTTTGCCAATGACG	798
	/dev_stage="adult"	QY	694	GCGCTGTGAGCTGACCTGCTCCAGGAAACCACTCAGCCACATCCAGCCCTTTTCTTTT	753
	/notes=Host: TOP10; Vector: pME18S-PL3 (Acc.No. AB009864);	Db	799	GCGCTGTGAGCTGACCTGCTCCAGGAAACCACTCAGCCACATCCAGCCCTTTTCTTTT	858
	R. Site1: DraIII (CACTGTGTG); R. Site2: DraIII	QY	754	CTGACCTTCGAGAGCTCCGCTCCCTGATCTTGACAGCAATCGCTGCCAAGCTTTGGG	813
	(CACCATGTG)	Db	859	CTGACCTTCGAGAGCTCCGCTCCCTGATCTTGACAGCAATCGCTGCCAAGCTTTGGG	918
	571. 2940	QY	814	GAGGACACCTTCGCGGGCTCGTCAACCTGCGAGCACTTATCTGTAACAAACACGAGCTG	873
	/codon_start=1	Db	919	GAGGACACCTTCGCGGGCTCGTCAACCTGCGAGCACTTATCTGTAACAAACACGAGCTG	978
	/product="hypothetical protein"	QY	874	GCGGCACTCGAGATGAGGCTTTTTCGAGCACTTCTGCTGACATTTGAGAGATCTGACCTC	933
	/protein_id="BAB39323.1"	Db	979	GCGGCACTCGAGATGAGGCTTTTTCGAGCACTTCTGCTGACATTTGAGAGATCTGACCTC	1038
	/db_xref="GI:13365899"	QY	934	TCCTTCAACAACTTCATGCTGCTGCGCTGCGGACTCCGTCGCGAGCATGCTCAACCTCCAC	993
	/translators="METLIGLLAFMAFAVVDACPKYCVCONLSGLTCLPSKGLL	Db	1039	TCCTTCAACAACTTCATGCTGCTGCGCTGCGGACTCCGTCGCGAGCATGCTCAACCTCCAC	1098
	FVPPDIDRTVEIRLGNFIHISRODFANMTGLVLTLSRNTISHIQFSPFLDLESL	QY	994	CAGCTGAGCTGAGCAACCACTGCTGATCACTCGCGGAGCACTTTTTCGAGACCTG	1053
	RSLLDNRSLPGEEDTLRGLVNLHIVNNQGLGADFEFDLTLEDLDSLNN	Db	1099	CAGCTGAGCTGAGCAACCACTGCTGATCACTCGCGGAGCACTTTTTCGAGACCTG	1158
	LHGLPMSVVRWNLHSLDHLHIAEFTADQLKARLDLTSNRLOKLPDPPIF	QY	1054	CAGAACTGCGCGCTGGATCTCACTCCAACTCGGCTGCAAGAGTCCGCTGATCC	1113
	ARSOALATPAPPLPSFSGNPLHNCCELLRLRLRDLDTGSGPGLKGRYF	Db	1159	CAGAACTGCGCGCTGGATCTCACTCCAACTCGGCTGCAAGAGTCCGCTGATCC	1218
	WHYREBFCEPPLIOTHKLLVLEGOATLCKAIGDPSPLIHWVAPDRLVGNS	QY	1114	ATCTTTTCCGCTCCAGGCTTCGGCTTTGACAGCACACCTTTTTCGCGGAGGCTC	1173
	RTAVYDNGTLDITTSQDSGATCTAANAAGTATVEVSIIVQLPHLSNSTRAPP	Db	1219	ATCTTTTCCGCTCCAGGCTTCGGCTTTGACAGCACACCTTTTTCGCGGAGGCTC	1278
	KSLRSLDITSGSGSGSGGEPKPPPERAVLSEVTTSALAKMSVSKSTPRVK	QY	1174	TTTGTAGTTTGGGGTAAACCACTTCACTGCAATTTGTAGCTTCTCTGGCTGGGAGGCTC	1233
	MYQLQNCSDDEVLVYRIPASNAKFAVNVLSGTGYDCLVLANWDDTATTLTATNIV	Db	1279	TTTGTAGTTTGGGGTAAACCACTTCACTGCAATTTGTAGCTTCTCTGGCTGGGAGGCTC	1338
	GCAQFTKADYPCQSHSILGTMILVIGIIVATLLVILVMRYKVNCHDSHS	QY	1234	GAGCGGAGCATGACCTGAAACCTGTGGTCTCCAGGGGCTCAAGGGTCTGCTACTTTC	1293
	KMAAASVNSQINGAPQPPPPAPAGAPQPPKVVVRNELDLFTASLARASDSSS	Db	1339	GAGCGGAGCATGACCTGAAACCTGTGGTCTCCAGGGGCTCAAGGGTCTGCTACTTTC	1398
	SSLGSGAAGLGRAPWRLPPSAPRPSLRLMGAFLASLDLKSQRKEELLDSTRPGR	QY	1294	TGSCATGTGCTGAGGAGGAGTTTGTGCGAGCCGCTCTCATCAACCCAGCAGACACAC	1353
	GAGTSARGHSRDREPLGPPAARASLLPLLEGKAKRSHSFDMGDFAAAAAGGVPG	Db	1399	TGSCATGTGCTGAGGAGGAGTTTGTGCGAGCCGCTCTCATCAACCCAGCAGACACAC	1458
	GYSPPRVSNWIKRSLVSNGMLLPFEESDLVGARTFGSSENVMESTV"	QY	1354	AAGTTGTGTTCTTGGAGGCCAGGCGGCACTCAAGTGCAAAAGCATTTGGGAGCCCC	1413
	BASE COUNT 628 a 1071 c 967 g 654 t	Db			
ORIGIN					
Query Match	93.1%; Score 2927; DB 9; Length 3320;				
Best Local Similarity	96.5%; Pred. No. 0;				
Matches 3068; Conservative	0; Mismatches 75; Indels 36; Gaps 6;				
QY	1 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACATACACACCCCGGAGGCTCG	60			
Db	80 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACATACACACCC--GCAGGCTCG	138			
QY	61 TCTGAACCTGAAGACACCCACATCCCAAGATGCCGAGGTTCTCGGAATGCCTGGGT	120			
Db	139 TCTGAACCTGAAGACACCCACATCCCAAGATGCCGAGGTTCTCGGAATGCCTGGGT	198			
QY	121 TCTTCGATCCGGAATCCTACCGGCATCCTCTAGGAGGAGATATTATTATTAT--TT	178			
Db	199 TCTTCGATCCGGAATCCTACCGGCATCCTCTAGGAGGATATTATTATTATTAT	258			
QY	179 TCTTTTAATCTGGAAGAGAGAGAAAGTTGTGCTTTTCCCGCTTCTTCTGCTTAAC	238			
Db	259 TCTTTTAATCTGGAAGAGAGAGAAAGTTGTGCTTTTCCCGCTTCTTCTGCTAAAT	318			
QY	239 GCCATGGATATAACTGAATAACGGCTCAGGCTTTCCCGCGTGGAGCTCCGAGGCCAC	298			
Db	319 GCCATGGATATAACTGAATAACGGCTCAGGCTTTCCCGCGTGGAGCTCCGAGGCCAC	378			

CDS

FEATURES

Source



2539	ATGGGGCCCTTCGCTCCCTGGAGCTCAAGAGTCAAGAAAGGAGGAGCTGCTCGACTCC	2598
2494	AGGACTCCAGCCGGAGAGGGGCTGGGACGTCCGGCCCGGGGCCA <sup>2</sup> CACTCGGACCCGAGAG	2553
2599	AGGACTCCAGCCGGAGAGGGGCTGGGACGTCA <sup>2</sup> CGCAGAGGGGCCACCACTCGGACCCGAGAG	2658
2554	CCACTGCTGGGGCCCTCTCGGCCCGGGCCAGAGAGCTGCTCCCTTCGCGTTCGCTGGAGGGC	2613
2659	CCACTGCTGGGGCGCCCGCGGGCCGAGAGCTGCTCCCTTCGCGTTCGCGTTCGAGGGC	2718
2614	AAGGCCAAACCGCAGCCACTCTTC <sup>2</sup> TGCACATGGGGGACTTTGCTGTCGGCGCGCGGGAGGG	2673
2719	AAGGCCAAACCGAGTCACTCTTC <sup>2</sup> CGACATGGGGACTTTGCGGCTCGCGCGCGGGAGGG	2778
2674	GTCTGCGCGGGCGGCTACAGTCTCTCTCGGAAGGTCTCGAAACATCTCGACGAAGCGCAGC	2733
2779	GTAGTGGCGGGCGGCTACAGTCC <sup>2</sup> CCCTCGGAGGGTCT <sup>2</sup> CGAAACATCTCGACGAAGCGCAGC	2838
2734	CTCTCTGTCAACGGCATGCTCTTGCCCTTT <sup>2</sup> GAGGAGAGTCA <sup>2</sup> CGCTGGTGGGGGCCCGGGGG	2793
2839	CTCTCTGTCAACGGCATGCTCTTGCCCTTT <sup>2</sup> CGAGGAGAGTCA <sup>2</sup> CTGTGTGGGGGCCCGAGGG	2898
2794	ACTTTTGGCAGCTCCGAATCGGTGATGGAGAGCACGGTCTAGTGTGGGGTGGGCATGCTC	2853
2899	ACTTTTGGCAGCTCGGAGTGGGTCA <sup>2</sup> TGGAGAGCACGGTCTAGTGTGGGGTGGGCATGCTC	2958
2854	CC <sup>2</sup> TTTCTGTGCGCAGGGTGGG---AGAAGGGGAAGAATCTCACTCGCAAGTCTTTCTG	2910
2959	CC <sup>2</sup> TTTCTGTGCGCGAGGGTGGGAGAAGAAGGGGAAGAATCTCACTGGCAAGTGTGTTGTG	3018
2911	GAGTTTCCATGTGTGATGTTTACATCCAGGGACAGT <sup>2</sup> TCGCTCCCTCTCAATGGCCCTCGT	2970
3019	GAGTTTCCATGTGTGATGTTTACATCCAGGGACAGT <sup>2</sup> TTTGTCTCCCTGTCAATGGCCCTCGT	3078
2971	GT-----CCCCCCTACCCGCAACCCACATCACTCC <sup>2</sup> CCACCA <sup>2</sup> CCCGCGGGGTGT	3029
3079	GTCCCCCCCCCCCCAACCCCCCAACCCCATACCTCC <sup>2</sup> CCACCA <sup>2</sup> CCCGCGGGGGTGT	3138
3026	GCTCAGGGAATGTGGACTCGCTCAAA <sup>2</sup> TGCCGGA <sup>2</sup> CTGAGCCCTCGAGTGT <sup>2</sup> TGGAAAGGCGA	3085
3139	GCTCAGGGAATCGGACTCGCTCAAA <sup>2</sup> TACCGA <sup>2</sup> CTGAGCCCTGAGTGT <sup>2</sup> TGGAAAGGCGA	3198
3086	GACTCCGCCCTTCTAATCAAAATGTAGCCTACAA <sup>2</sup> GCAAGCGGCTTTGGATGTCTTATG	3144
3199	GACTCCGCCCTTCTAATCAAAATGTAGCCTACAA <sup>2</sup> GCAAGCGGCTTTGGATGTCTTATG	3257

## RESULT 3

AL591063	100426 bp	DNA	linear	PRI 19-JUL-2001					
LOCUS	Human DNA sequence from clone RP11-535K1 on chromosome 6, complete sequence.								
DEFINITION									
ACCESSION	AL591063								
VERSION	AL591063.6	GI:15020551							
KEYWORDS	HTG.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
TITLE	1..(bases 1 to 100426)								
JOURNAL	Direct Submission								
	Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk								
	Clone requests: clonerequest@sanger.ac.uk								
COMMENT	On Jul 25, 2001 this sequence version replaced gi:14787623.								
	During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.								
	This sequence was finished as follows unless otherwise noted: all								

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep). This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>. RP11-535K1 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>. VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-535K1. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-121P10 is at 98427 in this sequence. The true right end of clone RP11-552E20 is at 2000 in this sequence.

## FEATURES

## Location/Qualifiers

```

source
1..100426
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-535K1"
/clone_lib="RP11-11.2"

repeat_region
1..1036
/note="L1M3 repeat: matches 4727. .5917 of consensus"

repeat_region
1030..1221
/note="L1 repeat: matches 4267. .4460 of consensus"

repeat_region
1222..1296
/note="Alu repeat: matches 1. .50 of consensus"

repeat_region
1297..3135
/note="L1 repeat: matches 2379. .4267 of consensus"

repeat_region
3134..3536
/note="L1M1 repeat: matches 5186. .5588 of consensus"

repeat_region
3537..3675
/note="FLAM A repeat: matches 1. .142 of consensus"

repeat_region
3677..4233
/note="L1M1 repeat: matches 5582. .6164 of consensus"

repeat_region
4523..5546
/note="L1M4 repeat: matches 2429. .3489 of consensus"

repeat_region
5510..5956
/note="L1M5 repeat: matches 4780. .5243 of consensus"

repeat_region
5997..6289
/note="L1M5 repeat: matches 102. .365 of consensus"

repeat_region
6303..7075
/note="L1M5 repeat: matches 5264. .6022 of consensus"

repeat_region
7078..7139
/note="MSTA repeat: matches 365. .426 of consensus"

repeat_region
7147..7205
/note="Alu repeat: matches 244. .302 of consensus"

repeat_region
7201..7553
/note="MSTA repeat: matches 1. .366 of consensus"

repeat_region
7555..7710
/note="L1M6 repeat: matches 6020. .6172 of consensus"

repeat_region
9233..9423
/note="MIR repeat: matches 22. .219 of consensus"

repeat_region
9441..9522
/note="41 copies 2 mer tt 63% conserved"

repeat_region
11776..11955
/note="MIR repeat: matches 59. .235 of consensus"

repeat_region
15525..15661
/note="L2 repeat: matches 2594. .2744 of consensus"

repeat_region
15937..16034
/note="MIR repeat: matches 60. .262 of consensus"

repeat_region
18523..18696
/note="MIR repeat: matches 78. .256 of consensus"
18748..19107
/note="MIR33 repeat: matches 3. .324 of consensus"
20177..20290
/note="3 copies 38 mer 76% conserved"
20885..20965
/note="MIR repeat: matches 35. .115 of consensus"
22278..22945
/note="CpG island"
/evidence-not_experimental
23265..23372
/note="18 copies 6 mer agagag 64% conserved"
23279..23372
/note="47 copies 2 mer ag 67% conserved"
23488..23613
/note="63 copies 2 mer ga 58% conserved"
23498..23611
/note="19 copies 6 mer gacaga 63% conserved"
23532..23603
/note="4 copies 18 mer 88% conserved"
23533..23604
/note="2 copies 36 mer 90% conserved"
24897..25274
/note="THL1B repeat: matches 1. .364 of consensus"
25405..25583
/note="MIR repeat: matches 10. .180 of consensus"
25906..26082
/note="MIR repeat: matches 14. .218 of consensus"
26339..26599
/note="MIR repeat: matches 2. .262 of consensus"
26933..27010
/note="13 copies 6 mer tgtgtg 92% conserved"
27107..27225
/note="MIR repeat: matches 120. .262 of consensus"
27481..27801
/note="L2 repeat: matches 1904. .2246 of consensus"
28341..28546
/note="MER20 repeat: matches 9. .218 of consensus"
28712..28786
/note="LTR33 repeat: matches 10. .87 of consensus"
29265..29354
/note="LTR33 repeat: matches 110. .198 of consensus"
29670..29983
/note="AluX repeat: matches 1. .312 of consensus"
30423..30512
/note="L2 repeat: matches 2602. .2695 of consensus"
30662..30966
/note="AluX repeat: matches 1. .305 of consensus"
31122..31173
/note="26 copies 2 mer tg 71% conserved"
31489..31573
/note="MLT1J repeat: matches 103. .190 of consensus"
33390..33444
/note="L2 repeat: matches 2651. .2705 of consensus"
32405..32618
/note="MIR repeat: matches 3. .249 of consensus"
32729..32977
/note="MIR repeat: matches 14. .262 of consensus"
33132..33310
/note="L2 repeat: matches 2254. .2449 of consensus"
33922..34526
/note="L2 repeat: matches 2059. .2672 of consensus"
34553..34647
/note="L2 repeat: matches 2658. .2749 of consensus"
34662..34869
/note="MER3 repeat: matches 1. .209 of consensus"
35504..35771
/note="MIR repeat: matches 2. .257 of consensus"
35720..35782
/note="L2 repeat: matches 2688. .2750 of consensus"
36322..36529
/note="MIR repeat: matches 22. .255 of consensus"
37340..37724

```

repeat_region	/note="MIR repeat: matches 4. .188 of consensus"	
repeat_region	38100. .38197	
repeat_region	/note="U6 repeat: matches 2. .103 of consensus"	
repeat_region	38228. .38496	
repeat_region	/note="L1MB3 repeat: matches 5903. .6181 of consensus"	
repeat_region	38815. .38998	
repeat_region	/note="MIR repeat: matches 8. .228 of consensus"	
repeat_region	39066. .39383	
repeat_region	/note="MSTD repeat: matches 1. .394 of consensus"	
repeat_region	39778. .39889	
repeat_region	/note="MIR repeat: matches 53. .165 of consensus"	
repeat_region	40464. .40529	
repeat_region	/note="MIR repeat: matches 51. .120 of consensus"	
repeat_region	40657. .40862	
repeat_region	/note="MIR repeat: matches 20. .249 of consensus"	
repeat_region	41569. .41640	
repeat_region	/note="36 copies 2 mer gt 66% conserved"	
repeat_region	41737. .42201	
repeat_region	/note="L2 repeat: matches 2203. .2701 of consensus"	
repeat_region	42225. .42370	
repeat_region	/note="MIR repeat: matches 116. .262 of consensus"	
repeat_region	42733. .42853	
repeat_region	/note="MIR repeat: matches 13. .128 of consensus"	
repeat_region	42907. .43213	
repeat_region	/note="L2 repeat: matches 2452. .2750 of consensus"	
repeat_region	44325. .44471	
repeat_region	/note="MIR repeat: matches 31. .184 of consensus"	
repeat_region	44544. .44770	
repeat_region	/note="MIR repeat: matches 27. .262 of consensus"	
repeat_region	45462. .45770	
Query Match 45.3%; Score 1424.8; DB 9; length 100426;		
Best Local Similarity 97.2%; Pred. No. 3.9e-237;		
Matches 1450; Conservative 0; Mismatches 42; Indels 0; Gaps 0;		
Qy	420	ACCGGGGTAGATGCTGCTGCCAGGCGTGAAGTACACAGACCATGAGACCCCTGCT 479
Db	63405	ACCCTGAGCTTGTGTTCTCCCTTCCAGGCGTGAAGTACACAGACCATGAGACCCCTGCT 63346
Qy	480	TGTTGGCTGCTAGCTTGGATGCGTTTGGCGTGGTGGCGCTGCCCAAGTACTG 539
Db	63345	TGTTGGCTGCTAGCTTGGATGCGTTTGGCGTGGTGGCGCTGCCCAAGTACTG 63286
Qy	540	TGTTGCCAGATCTGCTGAGTCACTGGGGACCTGTGCCCTCCAGGGGCTGCTTT 599
Db	63285	TGTTGCCAGATCTGCTGAGTCACTGGGGACCTGTGCCCTCCAGGGGCTGCTTT 63226
Qy	600	TGTACCCCTGATTTGACGGCGGACAGTGGAGCTGGCGCTGGGGGCAACTTCATAT 659
Db	63225	TGTACCCCTGATTTGACGGCGGACAGTGGAGCTGGCGCTGGGGGCAACTTCATAT 63166
Qy	660	CCACATCAGCGCCAGGACTTTGCCAATGACGGGGCTGGTGGACCTGACCCCTGTCCAG 719
Db	63165	CCACATCAGCGCCAGGACTTTGCCAATGACGGGGCTGGTGGACCTGACCCCTGTCCAG 63106
Qy	720	GAACACCATCAGCACATCCAGCCCTTTCTTTCTGGACCTCGAGAGCTCGCGTCCCT 779
Db	63105	GAACACCATCAGCACATCCAGCCCTTTCTTTCTGGACCTCGAGAGCTCGCGTCCCT 63046
Qy	780	GCATCTTGACGAATCGGCTGCCAAGCCTTGGGAGGACACCCCTCGGGGCTGGTCAA 839
Db	63045	GCATCTTGACGAATCGGCTGCCAAGCCTTGGGAGGACACCCCTCGGGGCTGGTCAA 62986
Qy	840	CCTGCAGCACCTTATCGTGAACAAACACAGCTGGGGGCTGCGAGATGAGCTTTTCA 899
Db	62985	CCTGCAGCACCTTATCGTGAACAAACACAGCTGGGGGCTGCGAGATGAGCTTTTCA 62926
Qy	900	GGATCTTCTGCTGACATTTGAGGATCTGGACCTCTCTTACAAACAACTCATGGGCTGCC 959
Db	62925	GGATCTTCTGCTGACATTTGAGGATCTGGACCTCTCTTACAAACAACTCATGGGCTGCC 62866
Qy	960	GTGGGATCTCGTGGACGATGCTCAACCTCCACAGCTGAGCTGGACCAACCTGCT 1019

RESULT 4  
AC011679  
LOCUS  
DEFINITION  
Homo sapiens clone RP11-14G23, WORKING DRAFT SEQUENCE, 16 unordered  
pieces.  
AC011679  
VERSION  
AC011679.3 GI:7329391  
KEYWORDS  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE  
Homo sapiens  
ORGANISM  
Homo sapiens

Db	62865	GTGGAGACTCCGTGCGACGACATGGTCAACCTCCACAGCTGAGCCCTGGACCAACACCTGCT 62806
Qy	1020	GGATCACATCGCGGAGGACACCTTTTGACAGACCTTGAGAACTGGCCCGCTGGATCTCAC 1079
Db	62805	GGATCACATCGCGGAGGACACCTTTTGACAGACCTTGAGAACTGGCCCGCTGGATCTCAC 62746
Qy	1080	CTCCAAATCGGCTGCGAGAGCTGCCCTTGATCCATCTTTGGCCGCTCCAGGCTTCGGC 1139
Db	62745	CTCCAAATCGGCTGCGAGAGCTGCCCTTGATCCATCTTTGGCCGCTCCAGGCTTCGGC 62686
Qy	1140	TTTGACAGCCACACCTTTGGCCCGCTTGTCTTGTAGTTTGGGGTAAACCACTTCA 1199
Db	62685	TTTGACAGCCACACCTTTGGCCCGCTTGTCTTGTAGTTTGGGGTAAACCACTTCA 62626
Qy	1200	CTGCAATTGTGAGCTTCTCTGCTCGGAGGCTCGAGCGGGACGATGACCTGGAACCTG 1259
Db	62625	CTGCAATTGTGAGCTTCTCTGCTCGGAGGCTCGAGCGGGACGATGACCTGGAACCTG 62566
Qy	1260	TGGTCCCGCAGGGGCTCAAGGTCGCTACTTCTGGCATGTGCGTGAAGAGTTTGT 1319
Db	62565	TGGTCCCGCAGGGGCTCAAGGTCGCTACTTCTGGCATGTGCGTGAAGAGTTTGT 62506
Qy	1320	GTCCGAGCGCTCTCATCACCCAGCACACACAGTTGCTGTTCTCGAGGGCCAGGC 1379
Db	62505	GTCCGAGCGCTCTCATCACCCAGCACACACAGTTGCTGTTCTCGAGGGCCAGGC 62446
Qy	1380	GGCCACACTCAAGTGCACAAAGCATTGGGACCCAGCCCTTATCCACTGGGTAGCCCC 1439
Db	62445	GGCCACACTCAAGTGCACAAAGCATTGGGACCCAGCCCTTATCCACTGGGTAGCCCC 62386
Qy	1440	CGATCAACCGCTGTGAGGAACTCTCAAGAGCCGCTGTCTATGACAATGGCACCTGGA 1499
Db	62385	CGATCAACCGCTGTGAGGAACTCTCAAGAGCCGCTGTCTATGACAATGGCACCTGGA 62326
Qy	1500	CATCTTCATCACACATCTCAGGACAGTGGTGGCTTCACTGCATGCTGCCAATGCTGC 1559
Db	62325	CATCTTCATCACACATCTCAGGACAGTGGTGGCTTCACTGCATGCTGCCAATGCTGC 62266
Qy	1560	CGGAGAGGCGACGGCATGGTGGAGGTCTCCATCTGCTCCAGCTGCCACACTCAGCAACAG 1619
Db	62265	CGGAGAGGCGACGGCATGGTGGAGGTCTCCATCTGCTCCAGCTGCCACACTCAGCAACAG 62206
Qy	1620	CACCAAGCGCTGACACCCCAAGTCCGCTCTCAGACATCACTGGCTCCAGCAAGAC 1679
Db	62205	CACCAAGCGCTGACACCCCAAGTCCGCTCTCAGACATCACTGGCTCCAGCAAGAC 62146
Qy	1680	CAGCCGGGAGGTGGAGGAGTGGGGGCGGAGAGCTCCCAAAGCCCGGACCGGC 1739
Db	62145	CAGCCGGGAGGTGGAGGAGTGGGGGCGGAGAGCTCCCAAAGCCCGGACCGGC 62086
Qy	1740	TGTGCTTGTCTGAAGTGACCAACACCTCGGCGCTGGTCAAGTGGTCTGTGACGAAGTC 1799
Db	62085	TGTGCTTGTCTGAAGTGACCAACACCTCGGCGCTGGTCAAGTGGTCTGTGACGAAGTC 62026
Qy	1800	AGCACCCCGGTGAAGATGTACAGCTGCAGTACAACTGCTCTGACGATGAGTACTGAT 1859
Db	62025	AGCACCCCGGTGAAGATGTACAGCTGCAGTACAACTGCTCTGACGATGAGTACTGAT 61966
Qy	1860	TTACAGGATGATCCAGCTCCAAAGGCTTCTGGTCAACAACTGCTG 1911
Db	61965	TTACAGGATGATCCAGCTTGTGGCAAGGCGAGGGGTGCTGGCTGGG 61914

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;  
Mammalia: Eutheria: Primates: Catarrhini, Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

1 (bases 1 to 179820)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens, clone RP11-14G23  
Unpublished

2. (bases 1 to 179820)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,  
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,  
Galan,J., Gardyna,S., Grant,G., Hago,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,K., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tefaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (10-OCT-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 26, 2000 this sequence version replaced gi:6249709.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WBIR  
Web site: http://www.seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
Project Information  
Center project name: L3400  
Center clone name: 14 G 23

Summary Statistics  
Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 166032 bases at least Q40  
Consensus quality: 172101 bases at least Q30  
Consensus quality: 174784 bases at least Q20  
Insert size: 182000; agarose-fp  
Insert size: 178320; sum-of-contigs  
Quality coverage: 4.1 in Q20 bases; agarose-fp  
Quality coverage: 4.2 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently  
consists of 16 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1960: contig of 1960 bp in length  
1961 2060: gap of 100 bp  
2061 3217: contig of 1157 bp in length  
3218 3317: gap of 100 bp  
3318 6117: contig of 2800 bp in length  
6118 6217: gap of 100 bp  
6218 8883: contig of 2666 bp in length  
8884 8983: gap of 100 bp  
8984 12533: contig of 3550 bp in length  
12534 12633: gap of 100 bp  
12634 17388: contig of 4755 bp in length  
17389 17488: gap of 100 bp  
17489 27094: contig of 9606 bp in length  
27095 27194: gap of 100 bp  
27195 38372: contig of 11178 bp in length  
38373 38472: gap of 100 bp  
38473 49446: contig of 10974 bp in length

49447 49546: gap of 100 bp  
49547 50664: contig of 11118 bp in length  
50665 60764: gap of 100 bp  
60765 74690: contig of 13926 bp in length  
74691 74790: gap of 100 bp  
74791 89780: contig of 14990 bp in length  
89781 89880: gap of 100 bp  
89881 109807: contig of 19927 bp in length  
109808 109907: gap of 100 bp  
109908 128003: contig of 18096 bp in length  
128004 128104: gap of 100 bp  
128105 148838: contig of 20735 bp in length  
148839 148939: gap of 100 bp  
148940 179820: contig of 30882 bp in length.

FEATURES

Location/Qualifiers  
1..179820  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-14G23"  
/clone\_lib="RPC1-11 Human Male BAC"  
1..1960  
/note="assembly\_fragment"  
2061..3217  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
3318..6117  
/note="assembly\_fragment"  
6218..8883  
/note="assembly\_fragment"  
8984..12533  
/note="assembly\_fragment"  
12634..17388  
/note="assembly\_fragment"  
17489..27094  
/note="assembly\_fragment"  
27195..38372  
/note="assembly\_fragment"  
38473..49446  
/note="assembly\_fragment"  
49547..60664  
/note="assembly\_fragment"  
60765..74690  
/note="assembly\_fragment"  
74791..89780  
/note="assembly\_fragment"  
89881..109807  
/note="assembly\_fragment"  
109908..128003  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right  
128104..148838  
/note="assembly\_fragment"  
148939..179820  
/note="assembly\_fragment"  
43444 a 43444 c 42485 g 45943 t 1504 others

BASE COUNT 46444 a 43444 c 42485 g 45943 t 1504 others  
ORIGIN

Query Match 40.8%; Score 1282; DB 2; Length 179820;  
Best Local Similarity 100.0%; Pred. No. 2e-212;  
Matches 1282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1863 CAGGATATCCAGCCTCCACCAAGGCCTTCGGTCAACAACTGGTGTCAAGGACTGG 1922  
DB 112240 CAGGATATCCAGCCTCCACCAAGGCCTTCGGTCAACAACTGGTGTCAAGGACTGG 112299  
QY 1923 CTACGACTTGTGTGTCTGGCCATGTGGATGACACGCCACCACTACGGCCACCA 1982  
DB 112300 CTACGACTTGTGTGTCTGGCCATGTGGATGACACGCCACCACTACGGCCACCA 112359  
QY 1983 CATCGTGGCTGCCGCCAGTTCCTCACCAGGCTGACTACCGCAGTCCAGTCCATGCA 2042

112360	DB	CATCGTGGGCTGGCGCCAGTTCTTTCACAAAGGCTGACTACCGCGAGTGCAGTGCATATGCA	112419
2043	QY	CAGCCAGATTCTGGGCGGCACCATGATCCTTGGTCAATCGGGGGGCATCATCTGTGGCCACGCT	2102
112420	DB	CAGCCAGATTCTGGGCGGCACCATGATCCTTGGTCAATCGGGGGGCATCATCTGTGGCCACGCT	112479
2103	QY	GCTGGTCTTCAATCGTCACTCTCATAGTGTGGCTTACAAGGTCTGCAACACAGAGGCCCCGAG	2162
112480	DB	GCTGGTCTTCAATCGTCACTCTCATAGTGTGGCTTACAAGGTCTGCAACACAGAGGCCCCGAG	112539
2163	QY	CAAGATGGCAGCGGCGGTGAGCAATGTGTACTCGCAGACCAACAGGCGGCCACGACACCGCC	2222
112540	DB	CAAGATGGCAGCGGCGGTGAGCAATGTGTACTCGCAGACCAACAGGCGGCCACGACACCGCC	112599
2223	QY	TCCAAGCAGCGCACACAGCCGGGGGCCCCCGCAGGCGGCCCGCGAAGGTGGTGGTCGCGAA	2282
112600	DB	TCCAAGCAGCGCACACAGCCGGGGGCCCCCGCAGGCGGCCCGCGAAGGTGGTGGTCGCGAA	112659
2283	QY	CGAGCTCTCTGACTTTCACGCGCAGCTGGCCGCGGCCAGTGACTCTCTTCTCTCAGGCTC	2342
112660	DB	CGAGCTCTCTGACTTTCACGCGCAGCTGGCCGCGGCCAGTGACTCTCTTCTCTCAGGCTC	112719
2343	QY	CCTGGGCAGTGGGAGGCTGGGGGCTGGGAGCGGCCCTCTGGAGGATCCACACCTCCGC	2402
112720	DB	CCTGGGCAGTGGGAGGCTGGGGGCTGGGAGCGGCCCTCTGGAGGATCCACACCTCCGC	112779
2403	QY	CCCCGCCCCAAGCCACGCTTGACCGCTGATGGGGGCTTCGCGCTCCCTGGAGCTCAA	2462
112780	DB	CCCCGCCCCAAGCCACGCTTGACCGCTGATGGGGGCTTCGCGCTCCCTGGAGCTCAA	112839
2463	QY	GAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCGGGAGAGGGGCTGGGAC	2522
112840	DB	GAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCGGGAGAGGGGCTGGGAC	112899
2523	QY	GTCCGCCGGGGGCCACACTCGEACCGAGAGCACTGCTGGGGCCCCCTCGCGGCCGGGC	2582
112900	DB	GTCCGCCGGGGGCCACACTCGEACCGAGAGCACTGCTGGGGCCCCCTCGCGGCCGGGC	112959
2583	QY	CAGGAGCCTGCTCCCTTCCGTTGGAGGGCAAGGCCAAACGACGCCACTCTCTTCGACAT	2642
112960	DB	CAGGAGCCTGCTCCCTTCCGTTGGAGGGCAAGGCCAAACGACGCCACTCTCTTCGACAT	113019
2643	QY	GGGGGACTTTCTGCTCGCGGGCGGGAGGGGTCTGTGCGGGCGGCTACAGTCTCTCTCG	2702
113020	DB	GGGGGACTTTCTGCTCGCGGGCGGGAGGGGTCTGTGCGGGCGGCTACAGTCTCTCTCG	113079
2703	QY	GAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGGATGCTCTTGCCCTT	2762
113080	DB	GAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGGATGCTCTTGCCCTT	113139
2763	QY	TGAGGAGTACCTGTGGGGGCCCGGGGACTTTTGGCAGCTCCGATGGGTGATGGA	2822
113140	DB	TGAGGAGTACCTGTGGGGGCCCGGGGACTTTTGGCAGCTCCGATGGGTGATGGA	113199
2823	QY	GAGCACGGTCTAGTGGGGGTGGGCATGCTCCCTTTCTGTGCGCAGGGTGGGAGAGGG	2882
113200	DB	GAGCACGGTCTAGTGGGGGTGGGCATGCTCCCTTTCTGTGCGCAGGGTGGGAGAGGG	113259
2883	QY	GAAGAATCTCACTGGCAAGTGTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGAC	2942
113260	DB	GAAGAATCTCACTGGCAAGTGTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGAC	113319
2943	QY	AGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCCTACCCCGCAACCCACACATCAC	3002
113320	DB	AGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCCTACCCCGCAACCCACACATCAC	113379
3003	QY	TCCCCACCAACCCGGCGGGGTGTGTCTCAGGGAATGTGGAATGTGGAATGTGGAATGTGGA	3062
113380	DB	TCCCCACCAACCCGGCGGGGTGTGTCTCAGGGAATGTGGAATGTGGAATGTGGAATGTGGA	113439
3063	QY	CCCTGTAGTGTGTGAAAGGGGAGACTCCGCTCTTCTAATCAAAATGTAGCTTACAAGC	3122
113440	DB	CCCTGTAGTGTGTGAAAGGGGAGACTCCGCTCTTCTAATCAAAATGTAGCTTACAAGC	113499

Qy 3123 AAGCGGCTTTGGATTGCTTATG 3144  
|||||  
Db 113500 AAGCGGCTTTGGATTGCTTATG 113521  
|||||

RESULT 5  
AC101792  
LOCUS  
DEFINITION AC101792 159443 bp DNA linear HTG 21-AUG-2002  
unordered pieces.  
AC101792  
VERSION AC101792.2 GI-22381528  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 159443)  
Birren,B., Nusbaum,C. and Lander,E.  
Mus musculus, clone RP24-176G22  
2 (bases 1 to 159443)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Govette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
McClean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,  
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 159443)  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 21, 2002 this sequence version replaced g1:17060567.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)

```

ftp://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WITR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: LI7407
Center clone name: 176_G22
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 152592 bases at least Q40
Consensus quality: 155716 bases at least Q30
Consensus quality: 156921 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 157543; sum-of-contigs
Quality coverage: 7.8 in Q20 bases; agarose-fp
Quality coverage: 7.8 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1129: contig of 1129 bp in length
1130 1239: gap of 100 bp
1230 2261: contig of 1032 bp in length
2262 2361: gap of 100 bp
2362 3703: contig of 1342 bp in length
3704 3803: gap of 100 bp
3804 5915: contig of 2112 bp in length
5916 6015: gap of 100 bp
6016 8269: contig of 2254 bp in length
8270 8369: gap of 100 bp
8370 9416: contig of 1047 bp in length
9417 9516: gap of 100 bp
9517 12017: contig of 2501 bp in length
12018 12117: gap of 100 bp
12118 13953: contig of 1836 bp in length
13954 14053: gap of 100 bp
14054 16946: contig of 2893 bp in length
16947 17046: gap of 100 bp
17047 20068: contig of 3022 bp in length
20069 20168: gap of 100 bp
20169 22610: contig of 2442 bp in length
22611 22710: gap of 100 bp
22711 25193: contig of 2483 bp in length
25194 25293: gap of 100 bp
25294 28313: contig of 3020 bp in length
28314 28413: gap of 100 bp
28414 31724: contig of 3311 bp in length
31725 31824: gap of 100 bp
31825 39332: contig of 8108 bp in length
39333 40032: gap of 100 bp
40033 50334: contig of 10302 bp in length
50335 50434: gap of 100 bp
50435 66250: contig of 15916 bp in length
66251 66350: gap of 100 bp
66351 78871: contig of 12521 bp in length
78872 78971: gap of 100 bp
78972 118433: contig of 39462 bp in length
118434 118533: gap of 100 bp
118534 159443: contig of 40910 bp in length.
Location/Qualifiers
1. .159443
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-176G22"/

```

FEATURES  
source

Db 92318 GTGAGCAGAGGCTACCGTGTGCTAGGTGGACAGGACGCT 92362

RESULT 6  
AC124583/C

LOCUS AC124583 303563 bp DNA linear HTG 25-JUN-2002  
DEFINITION Mus musculus chromosome UNK clone RP23-123P18, WORKING DRAFT  
SEQUENCE, 14 unordered pieces.  
AC124583  
AC124583.2 GI:21553290  
HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 303563)  
McPherson,J.D. and Waterston,R.H.  
TITLE The sequence of Mus musculus clone  
JOURNAL Unpublished  
2 (bases 1 to 303563)  
McPherson,J.D. and Waterston,R.H.  
REFERENCE Direct Submission  
3 (bases 1 to 303563)  
McPherson,J.D. and Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-JUN-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
On Jun 25, 2002 this sequence version replaced gi:21426723.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
----- Project Information -----  
Center project name: M.BA0123P18  
----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Sequencing vector: plasmid; 100%  
Chemistry: Dye-primer ET; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 297108 bases at least Q40  
Consensus quality: 298758 bases at least Q30  
Consensus quality: 299680 bases at least Q20  
Insert size: 211000; agarose-fp  
Insert size: 302263; sum-of-contigs  
Quality coverage: 0.00 in Q20 bases; agarose-fp  
Quality coverage: 10.12 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 14 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 1828: contig of 1828 bp in length  
\* 1829 1928: gap of unknown length  
\* 1829 3716: contig of 1788 bp in length  
\* 3717 3816: gap of unknown length  
\* 3817 5110: contig of 1294 bp in length  
\* 5111 5210: gap of unknown length  
\* 5211 9311: contig of 4101 bp in length  
\* 9312 9412: gap of unknown length  
\* 9412 15660: contig of 6249 bp in length  
\* 15661 15760: gap of unknown length  
\* 15761 21603: contig of 5843 bp in length  
\* 21604 21703: gap of unknown length



Tue Jun 10 08:57:36 2003

FEATURES	Source
21704	33356: contig of 11653 bp in length
33357	33456: gap of unknown length
33457	44470: contig of 11014 bp in length
44471	44570: gap of unknown length
44571	44570: gap of unknown length
44571	44571: contig of 32284 bp in length
44571	76854: contig of 32284 bp in length
76855	76934: gap of unknown length
76955	115914: contig of 38960 bp in length
115915	116014: gap of unknown length
116015	144525: contig of 28511 bp in length
144526	144625: gap of unknown length
144626	200283: contig of 55658 bp in length
200284	200383: gap of unknown length
200384	302035: contig of 101652 bp in length
302036	302135: gap of unknown length
302136	303563: contig of 1428 bp in length.
302136	303563: contig of 1428 bp in length.
Location/Qualifiers	
1..303563	
/organism="Mus musculus"	
/db_xref="taxon:10090"	
/chromosome="UNK"	
/clone="RP23-123P18"	
1..1828	
/note="assembly_name:Contig24"	
1929..3716	
/note="assembly_name:Contig25"	
3817..5110	
/note="assembly_name:Contig26"	
5211..9311	
/note="assembly_name:Contig27"	
9412..15660	
/note="assembly_name:Contig28"	
15761..21603	
/note="assembly_name:Contig29"	
21704..33356	
/note="assembly_name:Contig30"	
33457..44470	
/note="assembly_name:Contig31"	
44571..76854	
/note="assembly_name:Contig32"	
76955..115914	
/note="assembly_name:Contig33"	
116015..144525	
/note="assembly_name:Contig34"	
144626..200283	
/note="assembly_name:Contig35"	
200384..302035	
/note="assembly_name:Contig36"	
302136..303563	
/note="assembly_name:Contig9"	
BASE COUNT	81812 a 67733 c 68961 g 83749 t 1308 others
ORIGIN	
Query Match	37.1%; Score 1166.6; DB 2; Length 303563;
Best Local Similarity	86.6%; Pred. No. 1.9e-192;
Matches 1286; Conservative	0; Mismatches 199; Indels 0; Gaps 0;
QY	426 GGGTAGATGCTCGCTCGCCGAGGCGCTGAGTGACAGACCATGAGACCCCTGTTGTGG 485
Db	275430 GTGTCTCTTTCCAGGTGCAGAGGCTGAGTGACAGACCATGAGACCCCTGTTGTGG 275371
QY	486 CTTGTAGCTTTGGCATGGCTTTTCCGCTGTCAGCCCTGCGCCAGTACTGTGTCTG 545
Db	275370 GTGTGTGGCTTTTGGCATGGCTTTTGTGTGTGTGCTGCTGCTGCTGCTGCTG 275311
QY	546 CCAGAAATCTGTCTGAGTCACTGGGACCCCTGTGCCCTCCAGGGGCTCTTTGTACC 605
Db	275310 CCAGAAATCTGTCTGAGTCACTGGGACCCCTGTGCCCTCCAGGGGCTCTTTGTACC 275251
QY	606 CCCTGATATTGACCGCGGACAGTGGAGCTGGCGCTTGGCGGCAACTTCATCATCCAT 665
Db	275250 CCCTGATATTGACCGGNGACGTTGGAGTGGCTCTGGGTGGCAACTTCATCATCCAT 275191

275190 TGGCGGCAAGACTTTTGGCAACATGACAGGGCTGGTGGATCTGACTTTTTCAGGAACAC 275131  
Db |||||  
726 CATCAGCAGATCAGAGCCCTTTTCCTTTCTGGACCTCGAGAGCCCTCGGCTCCCTGCATCT 785  
QY |||||  
275130 CATCAGCAGATCAGAGCCCTTTCTCTTCTGGACCTTTGAGAGCCCTCGCTCCCTGCACCT 275071  
Db |||||  
786 TGACAGCAATCGGCTGCAAGCCTTGGGGAGGACACCTTCGGGGCTCGTCAACTGCA 845  
QY |||||  
275070 TGACAGCAACCGGCTACCTAGCCTTGGGGAGGACACACTCCGGGGTCTGGTCAACTGCA 275011  
Db |||||  
846 GCACCTTATCGTGAACAACAACACAGCTGGGGCGCATCGCAGATGAGGCTTTTGAGGACTT 905  
QY |||||  
275010 ACACCTTATTTGGAACAATAACAGCTGGGTGGCATCGCTGATGATGCTCTTTGAAGACTT 274951  
Db |||||  
906 CTGTGTCATTTGGAGGATCTGGACCTCTCTCAACAACCTCCATGGCTCGCGTGGGA 965  
QY |||||  
274950 CTGTGTCATTTAGAAGACCTTGGACCTATCTTCAACAACCTCCACCGACTTCCCTGGGA 274891  
Db |||||  
966 CTCGCTGCAACGATGTTCACTCCACAGCTCGAGCCTTGACACCAACCTGCTTGGATCA 1025  
QY |||||  
274890 TTCCGTACGGCGATGGTCAACTTCATCATAGCTGAGTCTGGACCACAATCTGCTGGACCA 274831  
Db |||||  
1026 CATCGCGAGGGCACCTTTTGCAGACTGAGAACTGCGCCGCTGGATCTCACTCCAA 1085  
QY |||||  
274830 CATTGTGAGGGCACTTTTGCAGACTGAGAACTGCGCCGCTGGATCTCACTCCAA 274771  
Db |||||  
1086 TCGGCTGCAAGCTCGCCCTGTATCCCATCTTTCCCGCTCCAGGCTTCTTGTCTAAC 1145  
QY |||||  
274770 TCGGCTGCAAGAACTTTCCCGGACCCCATCTTTCCCGCTCCAGGCTTCTTGTCTAAC 274711  
Db |||||  
1146 AGCCACACCTTTGCCCCACCTTGTCTTTAGTTTGGGGTAAACCACTTCACTGCAA 1205  
QY |||||  
274710 TGCCACGCCCTTTGCCCAACCCCTGTCTTTAGTTTGGAGGAAACCCGCTGCACTGCAA 274651  
Db |||||  
1206 TTGTGACTTCTCTGGCTGCGAGGCTCGAGCGGACGATGATCTGGAAACCTGTGGCTC 1265  
QY |||||  
274650 TTGTGAGCTTCTCTGGCTGCGAGGCTGGAGGGGATGATGACTTGGAGACCTGTGGATC 274591  
Db |||||  
1266 CCCAGGGGGCTCAAGGGTGCTACTTCTGGGATGTCGCTGAGGAGGAGTTTGTGTGCGA 1325  
QY |||||  
274590 CCCGGGAAGTCTCAAGGGTGCTACTTTTGGCATATTCGTGAGGAGGAGTTTGTGTGCA 274531  
Db |||||  
1326 GCGGCTCTCATCACCAGCACACACAAGTTGCTGTGTTCTGGAGGCGCAGCGGCCAC 1385  
QY |||||  
274530 GCGGCTCTCATCACCAGCACACACAAGTCTGTGTTCTGGAGGCGCAGCGGCCAC 274471  
Db |||||  
1386 ACTCAAGTGAAGGCATTTGGGACCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGA 1445  
QY |||||  
274470 TCTTAAGTGAAGGCATTTGGGACCCAGCCCCCTTGTATCCACTGGGTAGCCCCCGATGA 274411  
Db |||||  
1446 CCGCTCTGTAGGAACTCTCTAAGGACCGCTGTCTATGACAATGGACCTTGGACATCTT 1505  
QY |||||  
274410 CCGCTTGGTGGGAACTCTTCCAGGACTGCTATGACAATGGACCTTGGACATCTT 274351  
Db |||||  
1506 CATCACACATCTCAGGAAGTGTGCTTCACTGTCATTCCTGCAATGTGCGGAGA 1565  
QY |||||  
274350 AATTACCACTCTCAGGACAGTGGACCTTTTACCTGCTATCGGCAATGCAAGGAGA 274291  
Db |||||  
1566 GGGCAGGCAATGGTGGAGTCTCCATCGTCCAGTGGCAACCTCAGCAACAGCACAG 1625  
QY |||||  
274290 GGCTACGGCCACTGTAGAGTCTCCATTGTGAGTCTCCATTGTGAGTCTCCCAACCTTGTAAAGGAGCTAG 274231  
Db |||||  
1626 CCGCACTGCACCCCCAGTCCCGCTCTCAGACATCACTGGCTCCAGCAAGACAGCGG 1685  
QY |||||  
274230 CCGGATGGCAACCCCCAAAGTCTCGCTTTGGGACATCAGGTTCCAGCAAGACAGCGG 274171  
Db |||||  
1686 GGGAGGTGGAGGAGTGGGGCGGAGGCTCCCAAAGCCCGCGACGGGCTGTGCT 1745  
QY |||||  
274170 GGGAGGCGGAGTGTGGGCTGGGAGCTCCCAAGGAGCTCCCAAAGGAGCCCAAGAGGGCTGTGCT 274111  
Db |||||  
1746 TGTGTCTGAAGTACCAACCACTCGGCGCTTGGTCAAGTGGTCTGTGAGAAGTACGACC 1805  
QY |||||

Db 274110 TGTGTCGATGTTACACACATCTGCTCCCTGGTCAAGTGGTCTGTGTCAGCAAGTCAGCACC 274051  
Qy 1806 CCGGGTGAAGATGATACCACTGTCAGTACCACTGCTCTGACGATGAGGTACTGATTACAG 1865  
Db 274050 CAGGGTGAAGATGATACCACTGTCAGTACCACTGTTCTGATGATGAGGTCTGATCTACAG 273991  
Qy 1866 GATGATCCAGGCTCAACAGGCTTCTGTTGTCACACACCTGGT 1910  
Db 273990 GTGAGCCAGGCTACCGTGTGTCAGGTGGGACAGGACGCT 273946

RESULT 7  
AC125047/c  
LOCUS AC125047 308013 bp DNA linear HTG 29-JUL-2002  
DEFINITION Mus musculus chromosome UNK clone RP23-354G14, WORKING DRAFT  
SEQUENCE, 7 unordered pieces.  
ACCESSION AC125047  
VERSION AC125047.2 GI:22002214  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 308013)  
AUTHORS McPherson,J.D. and Waterston,R.H.  
TITLE The sequence of Mus musculus clone  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 308013)  
AUTHORS McPherson,J.D. and Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUN-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
REFERENCE 3 (bases 1 to 308013)  
AUTHORS McPherson,J.D. and Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (29-JUL-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
COMMENT On Jul 29, 2002 this sequence version replaced gi:21490485.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
Contact: [submissions@wustl.wustl.edu](mailto:submissions@wustl.wustl.edu)  
----- Project Information -----  
Center project name: M BA0354G14  
----- Summary Statistics -----

Sequencing vector: M13; 0%  
Sequencing method: plasmid; 100%  
Chemistry: Dye-primer ET; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 303857 bases at least Q40  
Consensus quality: 305426 bases at least Q30  
Consensus quality: 306387 bases at least Q20  
Insert size: 196000; agarose-fp  
Insert size: 307413; sum-of-contigs  
Quality coverage: 0.00 in Q20 bases; agarose-fp  
Quality coverage: 9.67 in Q20 bases; sum-of-contigs  
----- NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 12436: contig of 12436 bp in length  
\* 12437 12536: gap of unknown length  
\* 12537 40935: contig of 28299 bp in length  
\* 40936 40935: gap of unknown length

FEATURES

source

1. 308013  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-354G14"  
/chromosome="UNK"  
1. 12436  
/note="assembly\_name:Contig10"

misc\_feature 12537..40835  
/note="assembly\_name:Contig11"  
misc\_feature 40936..58095  
/note="assembly\_name:Contig12"  
misc\_feature 58196..142400  
/note="assembly\_name:Contig13"  
misc\_feature 142501..202912  
/note="assembly\_name:Contig14"  
misc\_feature 203013..306579  
/note="assembly\_name:Contig15"  
misc\_feature 306680..308013  
/note="assembly\_name:Contig18"

BASE COUNT 90720 a 61748 c 61427 g 93518 t 600 others  
ORIGIN

Query Match 37.1%; Score 1166.6; DB 2; Length 308013;  
Best Local Similarity 86.6%; Pred. No. 1.9e-192;  
Matches 1286; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

Qy 426 GGGTAGATGCTGCTCGCCAGGCGCTGAGTGACACAGACCATGGAGACCTGCTGGTGG 485  
Db 26361 GTGTCCTTCTTCCAGGTGCAGAGGCTGAGTGACACAGACCATGGAGACTCTGCTGGTGG 26302

Qy 486 CCGTGTAGCGTTGGCATGGGCTTTGGCTGTCAGCGCTGCCCGACCTGCCCAAGTACTGTGTCG 545  
Db 26301 GCTGCTGGCTTTGGCATGGGCTTTGGCTGTCAGCGCTGCCCGACCTGCCCAAGTACTGTGTCG 26242

Qy 546 CCAGAACTGTCTGAGTCACTGGGACCTGTGCGCCCTCCAAAGGGGCTCTCTTTGTACC 605  
Db 26241 CCAGAACTGTCTGAGTCACTGGGACCTGTGCGCCCTCCAAAGGGGCTCTCTTTGTGCC 26182

Qy 606 CCTGATATTGACCGGGGACAGTGGAGCTGCGCTGGGGCGGCAACTTCATCATCCACAT 665  
Db 26181 CCTGATATTGACCGGGGACAGTGGAGCTGCGCTGGGGCGGCAACTTCATCATCCACAT 26122

Qy 666 CAGCGCCAGGACTTTGCCACATACCGGGGCTGGGACCTGACCTGTCGAGGAACAC 725  
Db 26121 TGCGCGCCAGGACTTTGCCACATACCGGGGCTGGGACCTGACCTGTCGAGGAACAC 26062

Qy 726 CATCAGCCACATCCAGCCCTTTCTTCTTGGACCTCGAGAGCTCCCGCTCCCTGCACT 785  
Db 26061 CATCAGCCACATCCAGCCCTTTCTTCTTGGACCTCGAGAGCTCCCGCTCCCTGCACT 26002

Qy 786 TCACAGCAATCGGCTGCCAAGCCTTGGGAGAGACACCTCCGGGGCTGGTCAACCTGCA 845  
Db 26001 TCACAGCAATCGGCTGCCAAGCCTTGGGAGAGACACCTCCGGGGCTGGTCAACCTGCA 25942

Qy 846 GCACCTTATCGTGACACACACACCTGGGGGCTGGGAGCTGGGCTTTTGGAGACTT 905  
Db 25941 ACACCTTATGTGAACAATAACACCTGGGTGGGCTGGGAGCTTTTGAAGACTT 25882

Qy 906 CCTGCTGACATTGGAGGATCTGCACTCTCTTACAAACACCTCCATGGCGCTCCGCTGGGA 965  
Db 25881 CCTGCTGACATTGGAGGATCTGCACTCTCTTACAAACACCTCCATGGCGCTCCGCTGGGA 25822

Qy 966 CTCCGTGGACGATGGTCAACCTCCACAGCTGAGCCTGGACCAACAACTGCTGGATCA 1025

ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 197893)  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,  
Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,  
Bourka, J., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burckett, C., Surrall, K.L., Byrd, N.C.,  
Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.I., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
DeLanay, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karleson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louseged, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maneswari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okunolu, G.,  
Oggunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, F.,  
Rivers, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sison, I.,  
Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalon, B., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, C., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

TITLE  
JOURNAL

2 (bases 1 to 197893)  
Worley, K.C.  
Direct Submission  
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GXFO  
Center clone name: CH230-5008  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 143066 bases at least Q40  
Consensus quality: 151342 bases at least Q30  
Consensus quality: 157286 bases at least Q20  
-----

COMMENT

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 63 contigs. The true order of the pieces

Db	25821	TTCCGTACGCGCATGGTCAACCTCCATCAGCTGAGTCTGACCAACAATCTGCTGGACCA	25762
Qy	1026	CATCGCCGAGGACACCTTTGACAGACTTCGAGAAATCGCCCGCTGGATCTCACTCAA	1085
Db	25761	CATTGCTGAGGACACTTTTGCAGACTTCGAGAAATCGCCCGCTGGACCTCATCTCAA	25702
Qy	1086	TGCGCTGCAAGAGCTGCCCTCTGATGTCATCTTTCCCGCTCCAGGCTTCGGCTTTGAC	1145
Db	25701	TGCGCTGCAAGAACTTCCCGCGGACCCCATCTTTCCCGCTCCAGGCTTCCTGCTAAC	25642
Qy	1146	AGCCACACCTTTGCCCCACCTTGTCTTTAGTTTGGGGTAAACCCACTTCACTGCAA	1205
Db	25641	TGCACGCGCTTTGCCCCACCTTGTCTTTAGTTTGGGGTAAACCCACTTCACTGCAA	25582
Qy	1206	TTGTGAGCTTCTCTGCTCGGAGGCTCGAGCGGAGCATGACCTTGGAACTGTGGCTC	1265
Db	25581	TTGTGAGCTTCTCTGCTCGGAGGCTCGAGCGGAGCATGACCTTGGAGACTGTGGATC	25522
Qy	1266	CCGAGGGGCTCAAGGCTGCTACTTCTGGCATGTGCGTGAGGAGAGTTTGTGTGCA	1325
Db	25521	CCCGGAAGTCTCAAGGCTGCTACTTCTGGCATGTGCGTGAGGAGAGTTTGTGTGCA	25462
Qy	1326	GCCGCTCTCATCACCCAGCACACACAAAGTTGCTGTTCTGGAGGCGAGCGCCAC	1385
Db	25461	GCCGCTCTCATCACCCAGCACACACAAAGTTGCTGTTCTGGAGGCGAGCGCCAC	25402
Qy	1386	ACTCAAGTGAAAGCCATTGGGACCCAGCCCTTATCCATCTGCTGAGTACCCCTGATGA	1445
Db	25401	TCCTTAAGTGAAAGCCATTGGGACCCAGCCCTTATCCATCTGCTGAGTACCCCTGATGA	25342
Qy	1446	CGCTGCTAGGAACTCTCCAGGACTGCTGTATATGACAAATGGCACTTGGACATCT	1505
Db	25341	CGCTGCTAGGAACTCTCCAGGACTGCTGTATATGACAAATGGCACTTGGACATCT	25282
Qy	1506	CATCACCATCTCAGGACAGTGGTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTG	1565
Db	25281	AAATTACCACTCTCAGGACAGTGGTCTTCACTGCTGCTGCTGCTGCTGCTGCTG	25222
Qy	1566	GCCACGCGCATGTGGAGTCTCCATGCTCGAGCTGCGACCTCAGCAACAGCAGCAG	1625
Db	25221	GCTACGCGCATGTGGAGTCTCCATGCTCGAGCTGCGACCTCAGCAACAGCAGCAG	25162
Qy	1626	CCGCACTGACCCCAAGTCCCGCTCTCAGACATGCTGCTGCTGCTGCTGCTGCTG	1685
Db	25161	CCGCACTGACCCCAAGTCCCGCTCTCAGACATGCTGCTGCTGCTGCTGCTGCTG	25102
Qy	1686	GGGAGTGGAGGAGTGGGGCGGAGGCTCCCAAAAGCCCCCGGAAACGGGCTGTGCT	1745
Db	25101	GGGAGTGGAGGAGTGGGGCGGAGGCTCCCAAAAGCCCCCGGAAACGGGCTGTGCT	25042
Qy	1746	TGTGTCTGAAGTACCAACCACTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1805
Db	25041	TGTGTCTGAAGTACCAACCACTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	24982
Qy	1806	CCGGTGAAGATGATACAGTGCAGTACAACTGCTCTGACGATGAGTACTGATTACAG	1865
Db	24981	CAGGTTGAAGATGATACAGTGCAGTACAACTGCTCTGACGATGAGTACTGATTACAG	24922
Qy	1866	GATGATCCAGGCTCCCAAGGCTTCGTGGTCAACAACCTGGT	1910
Db	24921	GTGAGCCAGAGGCTACCGTGTGCTGAGTGGGACAGGACCGCT	24877

RESULT 8  
AC127880  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-5008, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
AC127880  
63 unordered pieces.  
AC127880  
VERSION  
AC127880.1 GI:21908317  
KEYWORDS  
HTG; HTGS PHASE1.  
SOURCE  
Rattus norvegicus.

\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1082: contig of 1082 bp in length  
\* 1083 1182: gap of unknown length  
\* 1183 2644: contig of 1462 bp in length  
\* 2645 2744: gap of unknown length  
\* 2745 3881: contig of 1137 bp in length  
\* 3882 3981: gap of unknown length  
\* 3982 5343: contig of 1362 bp in length  
\* 5344 6742: contig of 1299 bp in length  
\* 6743 6842: gap of unknown length  
\* 6843 8343: contig of 1501 bp in length  
\* 8344 8444: gap of unknown length  
\* 8444 10197: contig of 1753 bp in length  
\* 10197 10296: gap of unknown length  
\* 10297 11522: contig of 1226 bp in length  
\* 11523 11622: gap of unknown length  
\* 11623 13415: contig of 1793 bp in length  
\* 13416 13515: gap of unknown length  
\* 13516 15170: contig of 1655 bp in length  
\* 15171 15270: gap of unknown length  
\* 15271 17014: contig of 1744 bp in length  
\* 17015 17114: gap of unknown length  
\* 17115 18688: contig of 1574 bp in length  
\* 18689 18788: gap of unknown length  
\* 18789 20090: contig of 1302 bp in length  
\* 20091 20190: gap of unknown length  
\* 20191 22262: contig of 2072 bp in length  
\* 22263 22362: gap of unknown length  
\* 22363 24086: contig of 1724 bp in length  
\* 24087 24186: gap of unknown length  
\* 24187 25869: contig of 1683 bp in length  
\* 25870 25969: gap of unknown length  
\* 25970 28059: contig of 2090 bp in length  
\* 28060 28159: gap of unknown length  
\* 28160 30192: contig of 2033 bp in length  
\* 30193 30292: gap of unknown length  
\* 30293 32224: contig of 1932 bp in length  
\* 32225 32324: gap of unknown length  
\* 32325 34517: contig of 2193 bp in length  
\* 34518 34617: gap of unknown length  
\* 34618 35832: contig of 1215 bp in length  
\* 35833 35932: gap of unknown length  
\* 35933 38365: contig of 2433 bp in length  
\* 38366 38465: gap of unknown length  
\* 38466 39651: contig of 1186 bp in length  
\* 39652 39751: gap of unknown length  
\* 39752 42696: contig of 2945 bp in length  
\* 42697 42796: gap of unknown length  
\* 42797 44523: contig of 1727 bp in length  
\* 44524 44623: gap of unknown length  
\* 44624 47132: contig of 2509 bp in length  
\* 47133 47232: gap of unknown length  
\* 47233 49315: contig of 2083 bp in length  
\* 49316 49415: gap of unknown length  
\* 49416 52379: contig of 2964 bp in length  
\* 52380 52479: gap of unknown length  
\* 52480 54484: contig of 2005 bp in length  
\* 54485 54584: gap of unknown length  
\* 54585 57588: contig of 3004 bp in length  
\* 57589 59983: gap of unknown length  
\* 59984 60083: gap of unknown length  
\* 60084 62833: contig of 2750 bp in length  
\* 62834 62933: gap of unknown length  
\* 62934 65460: contig of 2527 bp in length  
\* 65461 67597: contig of 2037 bp in length  
\* 65561

\* 67598 67697: gap of unknown length  
\* 67698 71010: contig of 3313 bp in length  
\* 71011 73099: contig of 1989 bp in length  
\* 73100 73199: gap of unknown length  
\* 73200 76559: contig of 3460 bp in length  
\* 76560 76759: gap of unknown length  
\* 76760 79328: contig of 2569 bp in length  
\* 79329 79428: gap of unknown length  
\* 79429 82950: contig of 3522 bp in length  
\* 82951 83050: gap of unknown length  
\* 83051 85337: contig of 2287 bp in length  
\* 85338 85437: gap of unknown length  
\* 85438 88632: contig of 3195 bp in length  
\* 88633 88732: gap of unknown length  
\* 88733 92607: contig of 3875 bp in length  
\* 92608 92707: gap of unknown length  
\* 92709 95823: contig of 3116 bp in length  
\* 95824 95923: gap of unknown length  
\* 95924 98417: contig of 2494 bp in length  
\* 98418 98517: gap of unknown length  
\* 98519 102327: contig of 3710 bp in length  
\* 102328 102327: gap of unknown length  
\* 102329 106214: gap of unknown length  
\* 106215 106314: gap of unknown length  
\* 106315 109534: contig of 3220 bp in length  
\* 109535 109634: gap of unknown length  
\* 109635 113077: contig of 3443 bp in length  
\* 113078 113177: gap of unknown length  
\* 113178 116869: contig of 3692 bp in length  
\* 116870 116969: gap of unknown length  
\* 116970 121793: contig of 4824 bp in length  
\* 121794 121893: gap of unknown length  
\* 121894 126597: contig of 4704 bp in length  
\* 126598 126897: gap of unknown length  
\* 126898 131888: contig of 5191 bp in length  
\* 131889 131988: gap of unknown length  
\* 131989 136453: contig of 4465 bp in length  
\* 136454 136553: gap of unknown length  
\* 136554 141386: contig of 4833 bp in length  
\* 141387 141486: gap of unknown length  
\* 141487 146790: contig of 5304 bp in length  
\* 146791 146890: gap of unknown length  
\* 146891 151544: contig of 4654 bp in length  
\* 151545 151644: gap of unknown length  
\* 151645 157304: contig of 5660 bp in length  
\* 157305 157404: gap of unknown length

Query Match 36.6%; Score 1151.4; DB 2; Length 197893;  
Best Local Similarity 86.6%; Pred. No. 8.5e-190;  
Matches 1269; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 426 GGGTAGATGCTGCTCGCCAGCGCTGAGTCAGCAGACCATGGAGACCTGCTTGGTGG 485  
DB 42887 GTGTCTCTTCCAGTGCAGAGCGCGAGTCACAGACCATGGAGACTCTGCTTGGCGG 42946  
QY 486 CTTGCTAGCGTTTGGCATCGCTTTGCCGTGCTGAGCGCTGCCCAAGTACTGTGTCG 545  
DB 42947 GCTGCTGGCTTTTGGCATGGCGTTTGTGTTGTCGATGCTGCCCAAGTACTGTGTCG 43006  
QY 546 CCAGAAATCTGCTGAGTCACTGGGGACCTGTGTGCCCTCCAAAGGGGCTGCTCTTTGTACC 605  
DB 43007 CCAGAAATCTGCTGAGTCACTGGGACCTGTGTCTTCCAAAGGGGCTCTCTTGTGCC 43066  
QY 606 CCTGTATTTGACCGCGGACAGTGGAGCTGGCCCTGGGGGGAACCTTATCATCCACAT 665  
DB 43067 CCTGTATTTGACCGGAGGACAGTGGAGCTGGCTCTGGGTGGCAACTTATCATCCACAT 43126  
QY 666 CAGCCCCCAGGACTTTGCCCAACATGACGGGCTGTTGGACCTGACCTCTGTCCAGGAACAC 725  
DB 43127 TGGCCCCCAAGACTTTGCCCAACATGACAGGGCTAGTGGATCTGACCTTGTCTAGGAACAC 43186  
QY 726 CATCAGCCACATCCAGCCCTTTTCTCTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCT 785

Tue Jun 10 08:57:36 2003

43187	DB		CATCAGCCACATCCAGCCCTTCTCCTTCTCGACCTTGAGAGCTTCGGTTCCTCGACCT	43246
786	QY		TGACAGCANTCGCTGCCAAGCCTTGGGAGGACACCCCTCGGGGCTTGFTCAACCTGCA	845
43247	DB		TGACAGCAACCGGCTACCTAGCTTGGGAGGACACACTCCGGGCTTAGTCAACCTGCA	43306
846	QY		GCACCTTATCGTGAACAAACAACAGCTGGCGGATCGAGATGAGGCTTTTGAGGACTT	905
43307	DB		GCACCTTATGTGAACAATAATCAGCTGGGTGGCATTTGATGATGCTTTGAAGACTT	43366
906	QY		CTGCTGACATTCGAGGATCTGGACCTCTCTCAACAACCTCCATGGCTCCGTGGGA	965
43367	DB		CTGTGTGACTTTAGAGATTGGACCTATCTTACAACAACCTCCAGGCTTCCTCGGGA	43426
966	QY		CTCGTCGACGATGCTGAACCTCCACGCTGAGCTGGACCAACAACCTCTGGATCA	1025
43427	DB		TTCTGTAAAGCGCATGGTCAACCTCCATCAGCTGAGTCTGCACCAAACTTGTGGACCA	43486
1026	QY		CATCGCCGAGGACCTTTGAGACCTTCAGAAATCGGCCGCTGGATCACTCACTCCAA	1085
43487	DB		CATGCTGAGGTAACCTTCGACAGCTTCAGAAATCGGCCGCTGGACCTCAGTCCAA	43546
1086	QY		TCGGTGTGAGAAGTCCCTGATCCATCTTTGCCGCTCCAGGCTTCGGCTTTGAC	1145
43547	DB		TCGGTGTGAGAAGTTCCTCCCGAGACCACTTTGCCGCTCCAGGCTCTTGTCTCAC	43606
1146	QY		AGCCACACCTTTGCCCCACCTTGTCTCTTAGTTTGGGGTAAACCACTTCACCTGCAA	1205
43607	DB		TGCCACACCTTTGCCCCACCTGTCTCTTAGTTTGGGGTAAACCACTTCACCTGCAA	43666
1206	QY		TTGTGAGCTTCTTGCTCGCGAGGCTCGAGCGGACGATCAGCTTGGAACCTGTGGCT	1265
43667	DB		TTGTGAGCTTCTTGCTCGCGAGGCTCGAGCGGATCAGCAGCTGGAGACTGTGGATC	43726
1266	QY		CCAGGGGGCTCAAGGTCGTATCTTGTGCAATGTGCTGAGGAGAGTTTGTGTGCA	1325
43727	DB		CCCTGAGGTCCTCAAGGTCGTATTTTGGCATATTCGCGAGGAGAGTTTGTGTGA	43786
1326	QY		GCCTCTCTCATCACCAGCACACACAAGTTGCTGTTCTGAGGGCGAGCGGCCAC	1385
43787	DB		GCCGCCCTCATCACCAGCACACACAAGCTGCTGTTCTGAGGGCGAGCGGCCAC	43846
1386	QY		ACTCAAGTGCAAGCCATTGGGACCCGACCCCTTATCACTGGGTAGCCCCCGATGA	1445
43847	DB		TCCTAAGTGAAGGCCATTGGGACCCGACCCCTTATCACTGGTTGCCCCCGATGA	43906
1446	QY		CCGCTCGTAGGAACCTCTCAAGGACCGCTGTCTATGCAATGGCACCTCGACATTT	1505
43907	DB		CCGTTGGTAGGAACCTCTCCAGGACTGCGCTATATGACAACGGCACTCTGGACATTT	43966
1506	QY		CATCACACATCTCAGGACAGTGTGCTTCACTCGCATTTGCTGCCAATGCTCCGGAGA	1565
43967	DB		AATTAACCACTCTCAGGACGCGGCCCTTCACTGCTGATTCGAGCAACCGCGAGAGA	44026
1566	QY		GGCCACGCGCATGTGTGAGGCTCATGCTCCAGCTGCGACACTCAGCAACGACGACAG	1625
44027	DB		GGCTACCGCCACCGTAGAGGCTCTATGCTGACGCTCCACACTCAGTAAAGCACTAG	44086
1626	QY		CCGCACTGCAACCCCGAGTCCGCTCTCAGACATCACTGGCTCCAGCAAGACGACCGG	1685
44087	DB		CCGGATGGCGCGCCGACAGTCTGCTCTCAGACATCACTGGCTTCAGCAACGACGCTG	44146
1686	QY		GGGAGGTGAGGACGATGGGGGCGAGAGCTCCCAAGAGCCCCCGGAACGGGTGTGCT	1745
44147	DB		GGGAGGGGAGGTAGCGGGCTGGGAGGCTCCCAAGAGACCCACAGAGGGGTGTGCT	44206
1746	QY		TGTGTCTGAGTGACCAACCACTCGGCCCTGTGCAAGTGGTGTGTGACGAGTCAAGC	1805
44207	DB		TGTGTCTGATGTGACCAACCACTCGGCCCTGTGCAAGTGGTGTGTGACGAGTCAAGT	44266
1806	QY		CCGGGTGAAGATGATCAGGTGAGTACACTGCTCTGACGATGAGGTACTGATTTACAG	1865

Db 44267 CAGGGTGAAGATGATCAGCTGCAGTCACACTGTCCGACGACGAGGAGGATCGAATCGATTTTGCTG

QY 1866 GATGATCCAGCCCTCCACAAAGGCC 1890 HTG 19-JUL-2002  
LOCUS Rattus norvegicus clone CH230-5008, \*\*\* SEQUENCING IN PROGRESS \*\*\*,  
DEFINITION 63 unordered pieces.

Db 44327 GTGACCCAGAGCGCTGCAGTGTGGTC 44351

RESULT 9  
AC127880 197893 bp DNA linear HTG 19-JUL-2002  
Rattus norvegicus clone CH230-5008, \*\*\* SEQUENCING IN PROGRESS \*\*\*,  
LOCUS 63 unordered pieces.  
DEFINITION 63 unordered pieces.

ACCESSION AC127880 GI:21908317  
VERSION HTG: HTGS PHASE1.  
KEYWORDS Rattus norvegicus.  
SOURCE Rattus norvegicus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 197893)  
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amarutunge,H.C., Are,J.R., Ayelle,M., Banks,T.,  
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleaveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homai,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlasson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li/Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseg,H.,  
Lozadó,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E.I., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,N., Morgan,M., Morris,S.,  
Mosser,M., Neal,D., Newton,J., Newton,K., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokoko,S., Ogih,M., Okwuonu,G.,  
Oraguene,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telirod,B., Thomas,N., Thomas,S.,  
Umanai,K.I., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.

TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 197893)  
AUTHORS Worley,K.C.  
TITLE Submitted (19-JUL-2002) Human Genome Sequencing Center, Department  
JOURNAL of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
COMMENT ----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
WWW: http://www.hscg.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GXPO  
 Center clone name: CH230-5008  
 ----- Summary Statistics  
 Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 143066 bases at least Q40  
 Consensus quality: 151342 bases at least Q30  
 Consensus quality: 157286 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 63 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1082: contig of 1082 bp in length  
 \* 1083 1182: gap of unknown length  
 \* 1183 2644: contig of 1462 bp in length  
 \* 2645 2744: gap of unknown length  
 \* 2745 3881: contig of 1137 bp in length  
 \* 3882 3981: gap of unknown length  
 \* 3982 5343: contig of 1362 bp in length  
 \* 5344 5443: gap of unknown length  
 \* 5444 6742: contig of 1299 bp in length  
 \* 6743 6842: gap of unknown length  
 \* 6843 8343: contig of 1501 bp in length  
 \* 8344 8443: gap of unknown length  
 \* 8444 10196: contig of 1753 bp in length  
 \* 10197 10296: gap of unknown length  
 \* 10297 11522: contig of 1226 bp in length  
 \* 11523 11623: gap of unknown length  
 \* 11623 13415: contig of 1793 bp in length  
 \* 13416 13515: gap of unknown length  
 \* 13516 15170: contig of 1655 bp in length  
 \* 15171 15270: gap of unknown length  
 \* 15271 17014: contig of 1744 bp in length  
 \* 17015 17114: gap of unknown length  
 \* 17115 18688: contig of 1574 bp in length  
 \* 18689 18788: gap of unknown length  
 \* 18789 20090: contig of 1302 bp in length  
 \* 20091 20190: gap of unknown length  
 \* 20191 22262: contig of 2072 bp in length  
 \* 22263 22362: gap of unknown length  
 \* 22363 24086: contig of 1724 bp in length  
 \* 24087 24186: gap of unknown length  
 \* 24187 25869: contig of 1683 bp in length  
 \* 25870 25969: gap of unknown length  
 \* 25970 28059: contig of 2090 bp in length  
 \* 28060 28159: gap of unknown length  
 \* 28160 30192: contig of 2033 bp in length  
 \* 30193 30292: gap of unknown length  
 \* 30293 32224: contig of 1932 bp in length  
 \* 32225 32344: gap of unknown length  
 \* 32325 34517: contig of 2193 bp in length  
 \* 34518 34617: gap of unknown length  
 \* 34618 35832: contig of 1215 bp in length  
 \* 35833 35932: gap of unknown length  
 \* 35933 38355: contig of 2433 bp in length  
 \* 38356 38466: gap of unknown length  
 \* 38466 39651: contig of 1186 bp in length  
 \* 39652 39751: gap of unknown length  
 \* 39752 42696: contig of 2945 bp in length  
 \* 42697 42796: gap of unknown length  
 \* 42797 44523: contig of 1727 bp in length  
 \* 44524 44623: gap of unknown length  
 \* 44624 47132: contig of 2509 bp in length

\* 47133 47232: gap of unknown length  
 \* 47233 49315: contig of 2083 bp in length  
 \* 49316 49415: gap of unknown length  
 \* 49416 52379: contig of 2964 bp in length  
 \* 52380 52479: gap of unknown length  
 \* 52480 54484: contig of 2005 bp in length  
 \* 54485 54584: gap of unknown length  
 \* 54585 57588: contig of 3004 bp in length  
 \* 57589 59583: contig of 2295 bp in length  
 \* 59584 60083: gap of unknown length  
 \* 60084 62833: contig of 2750 bp in length  
 \* 62834 62934: gap of unknown length  
 \* 62934 65461: contig of 2527 bp in length  
 \* 65461 65560: gap of unknown length  
 \* 65561 67597: contig of 2037 bp in length  
 \* 67598 67697: gap of unknown length  
 \* 67698 71010: contig of 3313 bp in length  
 \* 71011 71110: gap of unknown length  
 \* 71111 73099: contig of 1989 bp in length  
 \* 73100 73199: gap of unknown length  
 \* 73200 76559: contig of 3460 bp in length  
 \* 76560 76660: gap of unknown length  
 \* 76660 79328: contig of 2569 bp in length  
 \* 79329 79428: gap of unknown length  
 \* 79429 82950: contig of 3522 bp in length  
 \* 82951 83050: gap of unknown length  
 \* 83051 85337: contig of 2287 bp in length  
 \* 85338 85437: gap of unknown length  
 \* 85438 88632: contig of 3195 bp in length  
 \* 88633 88732: gap of unknown length  
 \* 88733 92607: contig of 3875 bp in length  
 \* 92608 92707: gap of unknown length  
 \* 92708 95823: contig of 3116 bp in length  
 \* 95824 95923: gap of unknown length  
 \* 95924 98417: contig of 2494 bp in length  
 \* 98418 98518: gap of unknown length  
 \* 98518 102227: contig of 3710 bp in length  
 \* 102228 102327: gap of unknown length  
 \* 102328 106214: contig of 3887 bp in length  
 \* 106215 106314: gap of unknown length  
 \* 106315 109535: contig of 3220 bp in length  
 \* 109536 109634: gap of unknown length  
 \* 109635 113077: contig of 3443 bp in length  
 \* 113078 113177: gap of unknown length  
 \* 113178 116869: contig of 3692 bp in length  
 \* 116870 116969: gap of unknown length  
 \* 116970 121793: contig of 4824 bp in length  
 \* 121794 121894: gap of unknown length  
 \* 121895 126597: contig of 4704 bp in length  
 \* 126598 126697: gap of unknown length  
 \* 126698 131888: contig of 5191 bp in length  
 \* 131889 131988: gap of unknown length  
 \* 131989 136453: contig of 4465 bp in length  
 \* 136454 136553: gap of unknown length  
 \* 136554 141386: contig of 4833 bp in length  
 \* 141387 141486: gap of unknown length  
 \* 141487 146790: contig of 5304 bp in length  
 \* 146791 146890: gap of unknown length  
 \* 146891 151544: contig of 4654 bp in length  
 \* 151545 151644: gap of unknown length  
 \* 151645 157304: contig of 5660 bp in length  
 \* 157305 157404: gap of unknown length

Query Match 26.3%; Score 827.4; DB 2; Length 197893;  
 Best Local Similarity 80.9%; Pred. No. 1.1e-133;  
 Matches 1062; Conservative 0; Mismatches 216; Indels 35; Gaps 7;

QY 1859 TTTACAGATGATCCAGCCTCCAAACAGCCCTCGTGTGTCACACACCTGGTGTGTCAGGGA 1918

Db 194803 TCTGAGGATGATCCAGCCTCCAAACAGCCCTCGTGTGTCACACACCTGGTGTGTCAGGAA 194744

QY 1919 CTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACACCCACGACACTCAGCGCCA 1978





```
QY 1047 AGACCTGCGAAGAACTGGCCGCGCTGATCTCACTCAATCGGCTGCAAGAGCTGCCCC 1106
D 1106
D 603 CGCCTTGCAAGAGTGGCCGCGCTGACATGACTCCAAACCGCTGACCAACAATCCACC 662
QY 1107 TGATCCCATCTTTGGCCGCTCCAGGCTTCGGCTTTGAGAGCCACACCTTTTGCCCCACC 1166
D 1166
D 663 CGACCCACTCTTTCCCGGCTGCGCTGCTGCGAGGCGCGCGGCTGCGCCGCTCTGC 722
QY 1167 CTTGTCCTTTAGTTTGGGGTAACCCACTTCACTGCAATTTGAGCTTCTCTGGCTGG 1226
D 1226
D 723 CTTGGTCTGGCTTTGGCGGNAACCCCTGCACTGCAACTGCGAGCTGGTGGCTGG 782
QY 1227 GAGGCTGAGCGGAGCATGACTGGAACCTGTGGCTCCCGAGGGGCTCAAGGGTGG 1286
D 1286
D 783 TCGCCTGGCGGAGGAGCACTCGAGGCTCGCGTCCCACTGCTCGGGGCGG 842
QY 1287 CTACTCTGGCATGCTGGCTGAGGAGAGTTTGTGCGAGCGGCTCTCATCACCAGCA 1346
D 1346
D 843 CTACTCTGGCGGCTGGCGAGGAGAGTTTGTGCGAGCGGCTGCTGACTCAGCG 902
QY 1347 CACACAAAGTTGCTGTTCTGAGGCGCGAGGCGCACACTCAAGTGCAAGGCAATTGG 1406
D 1406
D 903 CTCACCACTCTGGCTGTCGCGCAGCTCGCGGCTGCCCTGCGCTGCGGCGAGTGG 962
QY 1407 GGACCCAGCGCCCTTATCACTGGGTAGCCCGGATGACCGCTGGTAGGAACCTCTC 1466
D 1466
D 963 GGACCCAGAGCGCGTGTGGTGTGCGTGTGTCAGGCGCGCTGCTAGGCACTCAAG 1022
QY 1467 AAGGACCGTGTATGACAAATGGCACCTGGACATCTTCATCACCACATCTCAGACAG 1526
D 1526
D 1023 CCGTGCCCGCGCTTCCCAATGGGACGCTGGTGTGCTGCTGCTGCTGCTGCTGCTG 1082
QY 1527 TGTGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1586
D 1586
D 1083 TGGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1142
QY 1587 CTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
D 1644
D 1143 GACTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1202
QY 1645 TCCCGCTCTGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1704
D 1704
D 1203 GAGCGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1262
QY 1705 GCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764
D 1764
D 1263 CGACACTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1314
QY 1765 ACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824
D 1824
D 1315 ACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1374
QY 1825 CTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1884
D 1884
D 1375 ATCCAGTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1434
QY 1885 AAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
D 1944
D 1435 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1494
QY 1945 ATGTGGGATGACACAGCCAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2004
D 2004
D 1495 GTGTATGAGACAGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1554
QY 2005 TTCACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2064
D 2064
D 1555 TCCACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1614
QY 2065 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2124
D 2124
D 1615 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1674
```

```
QY 2125 ATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2178
D 2178
D 1675 CTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1734
QY 2179 GTGACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2238
D 2238
D 1735 GTTAGCAGGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1794
QY 2239 GCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2257
D 2257
D 1795 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1813

RESULT 11
LOCUS BC003578 2549 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, clone MGC:2656 IMAGE:3537259, mRNA, complete cds.
ACCESSION BC003578
VERSION BC003578.1 GI:13097761
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2549)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Varon Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 8 Row: f Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomScan gene prediction..
FEATURES
source
1..2549
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:2656 IMAGE:3537259"
/tissue_type="lung, small cell carcinoma"
/clone_lib="NIH MGC 7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
453..2339
/codon_start=1
/product="Unknown (protein for MGC:2656)"
/protein_id="AAH03578.1"
/db_xref="GI:13097762"
/translation="MAILPLLCPLPLAPASSPPOSATPSPCPRCRCOTOSLPISVL
CPGAGLLVPFSLDRRAELADNFIAVRRRLDANMTGLLHLSLRSNTTRHVAAGA
FADURLALHLDGNRLTSLGEGQRLGNLRLHLLSNNAALAAAGALDCAETLED
LDLSYNNLEQLPWEALGRNVTGLDNLHLASVPAGAFSLRHLKLARLDMTSRLTT
```

BASE COUNT		440 a	935 c	738 g	436 t
ORIGIN					
Query Match 20.2%; Score 634.6; DB 9; Length 2549;					
Best Local Similarity 61.3%; Pred. No. 4e-100;					
Matches 1079; Conservative 0; Mismatches 664; Indels 16; Gaps 3;					
QY	507	GTITGCGGTGTCAGCGCTGCCCAAGTACTGTCTGCCAGAACTCTCTGAGTCACT	566		
DB	515	GTGAGCCACACCCAGGCCATGTCCCGCGCTGCCCTGCCAGACACAGTCTGCCCT	574		
QY	567	GCGGACCTGTGCCCTCCAAAGGGGCTCTTTGTAGCCCCGTGATATTGACGGGGAC	626		
DB	575	AAGCGTCTGTCCAGGGCAGCGCTCTGTGTCGCCACCTCTGCTGACCGCGGC	634		
QY	627	AGTGGAGCTGGCTGGCGGCACTTCATCCATCAGCCAGCGGAGCTTTGCCAA	686		
DB	635	AGCGAGCTGGCTGGCAGACAACTTCATCGCTCTCGTGGCGCGCGGACCTGCCAA	694		
QY	687	CATGACGGGCTGTGTGACCTGACCTCTGTCAGGAACACCATCAGCCATCCAGCCCTT	746		
DB	695	CATGACGGCTGTGATCTGAGCTGTGCGGGAACACCATCCGCGACGTGCTGCCG	754		
QY	747	TTCCCTTTTGAGCTCGAGAGCTCCGCTCCCTGATCTTGACAGCAATCGGCTGCCAAG	806		
DB	755	CGCTTCCGCGACCTGGCGGCGCTGCTGCTGCTGACCTGGATGGCAACCGGCTGACCTC	814		
QY	807	CTTTGGGAGACACCTCGGGGCTGTGCAACCTGCAGACCTTATCTGTAACAAACA	866		
DB	815	ACTGGCGAGGGCAGCTGCGCGGCTGTGCACTTGGCGCACTCTCATCTCAGCAACA	874		
QY	867	CCAGCTGGCGGCTCGCAGATAGGCTTTGAGGATCTCTGCTGATGAGGATCT	926		
DB	875	CCAGCTGACGCTGGCGGCGCGCTGATGATGTTGCGGAGACATCGGAGGACCT	934		
QY	927	GGACCTCTCTACAAACCTCCATGCGCTGCGTGGGACTCTCGTGGAGCGCATGGTCAA	986		
DB	935	CGACCTCTCTACAAACCTCGAGAGCTCCCTTGGAGGCTCTGGCGGCTGGGCAA	994		
QY	987	CTCCACCAAGCTGAGCTGGACCAACACCTGCTGATCACAATCGCGAGGCACTTTTC	1046		
DB	995	CGTCAACAGTTGGGCTCGACCAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1054		
QY	1047	AGACTGAGAACTGGCGGCTGGATCTACCTCCATCGGCTGCGAGAGCTGCCCC	1106		
DB	1055	CGGCTGCAACAGCTGGCGGCTGGATGATGATGATGATGATGATGATGATGATGAT	1114		
QY	1107	TGATCCCATCTTTGCGGCTCCAGGCTTGGGCTTTGACAGCAACACCTTTGCCAC	1166		
DB	1115	CGACCACTCTTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1174		
QY	1167	CTTTGCTCTTTAGTTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTTGCTGCG	1226		
DB	1175	CTTGGTGTGCTTTGGCGGGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1234		
QY	1227	GAGGCTGAGCGGAGATGACTTGAAACCTGTGCTCCCGAGGGGCTCAAGGGTCTG	1286		
DB	1235	TGCGCTGGCGGAGGAGGAGCTCTGAGGCTTGGCTCCCACTGCTCTGGCGGCG	1294		
QY	1287	CTACTTCTGGCATGTGCTGAGGAGGATTTGTGTCGAGCGGCTCTCATCACCAGCA	1346		
DB	1295	CTACTTCTGGCGGTGGCGGAGGAGTTTGTCTCGAGCGGCGGCTGCTGCTGCTGCT	1354		
QY	1347	CACACAAAGTTGCTGTTCTGAGGGCGAGGGCCACATCAAGTGCAAGAGCATTTGG	1406		

DB	1355	CTCACACACCTCTGGTGTGCTCCCGCAGGTGCGCGGCTGCCCTGCGCTGCGGCGAGTGGG	1414		
QY	1407	GGACCCAGGCCCCCTTATCACTGGGTAGCTACCCCGCATGACCGCTGGTAGGAACTCTTC	1466		
DB	1415	GGACCCAGGCCCCCTGTTGGGTGTACCCCGAGGCGGCTGCTAGGCAACTCAAG	1474		
QY	1467	AAGGACCGCTGTCTATGACAATGGCAACCTTGGACATCTTTCATCACACATCTCAGACAG	1526		
DB	1475	CGTGGCGGCGCTTCCCCAATGGAGCGTGGAGTGTCTGCTCACCAGCGGCTGATGG	1534		
QY	1527	TGTCGCTTCACTGATTTGCTGCCAATGCTGCGGAGAGGCGCACGCGCATGTGTGAGGT	1586		
DB	1535	TGGCATCTTCACTGATTTGCGGCAATGAGTGGGAGGCGCACAGCTGCTGTGTGAGCT	1594		
QY	1587	CTCATCGTCCAGTGCACACCTCAGCAACACAGCACAGCGCA--CTGACCCCCCAAG	1644		
DB	1595	GACTGTGGTCCCCCACCCTCTCAGCTAGCAACAGCACAGCTGTACCCCCCGCG	1654		
QY	1645	TCCGCGCTCTCAGACATCACTGGCTCCAGCAAGACAGCGGGAGGTGAGGACGTGGG	1704		
DB	1655	GGACGGGATCTGATGCTCTCACCCCACTGCTGCTGCTGCTGCTTCTGCCAAGGTGGC	1714		
QY	1705	GGCGAGAGCTCCAAAAGCCCCCGGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1764		
DB	1715	CGACACTGGGCCCCCTA-----CCGACGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCT	1766		
QY	1765	ACCTCGGCGCTGCTCAAGTGTGTGTGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	1824		
DB	1767	ACAGTGTCTTGTCCAGTGGCGGATCAGCGGCTATCCGCGCATCCGATGATGATGATGAT	1826		
QY	1825	CTGAGTACAATGCTGTGAGCATGAGTACTGATTTACAGGATGATCCAGCTCTCAAC	1884		
DB	1827	ATCCAGTACAACAGCTGGCTGATGATCTCTGCTTACAGATGATCCGCGCGAGAGC	1886		
QY	1885	AAGGCTTCTGCTGCTCAACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1944		
DB	1887	CGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1946		
QY	1945	ATGTCGGATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	2004		
DB	1947	GTGATGAGGACAGCGCCACGCGGCTCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2006		
QY	2005	TTACCAAGGCTGACTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2064		
DB	2007	TCACCGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2066		
QY	2065	ATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2124		
DB	2067	ATGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2126		
QY	2125	ATGTCGCTTACAAGTCTGCAACACAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2178		
DB	2127	CTAATCGCTTACAAGTCTGCAACAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2186		
QY	2179	GTGAGCAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2238		
DB	2187	GTTAGCAGGTTGCTTCCAGACACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2246		
QY	2239	GCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2257		
DB	2247	CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2265		

RESULT 12	AX266997	LOCUS	AX266997	Sequence 3	1887 bp	DNA	linear	PAT 26-OCT-2001
DEFINITION	AX266997	LOCUS	AX266997	Sequence 3	1887 bp	DNA	linear	PAT 26-OCT-2001
ACCESSION	AX266997	LOCUS	AX266997	Sequence 3	1887 bp	DNA	linear	PAT 26-OCT-2001
VERSION	AX266997.1	LOCUS	AX266997	Sequence 3	1887 bp	DNA	linear	PAT 26-OCT-2001
KEYWORDS	human.	LOCUS	AX266997	Sequence 3	1887 bp	DNA	linear	PAT 26-OCT-2001
SOURCE	human.	LOCUS	AX266997	Sequence 3	1887 bp	DNA	linear	PAT 26-OCT-2001
ORGANISM	Homo sapiens	LOCUS	AX266997	Sequence 3	1887 bp	DNA	linear	PAT 26-OCT-2001
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	LOCUS	AX266997	Sequence 3	1887 bp	DNA	linear	PAT 26-OCT-2001



Tue Jun 10 08:57:36 2003

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Glucksmann, M.A.  
 TITLE 33395, a novel human leucine-rich repeat family member and use  
 thereof  
 JOURNAL Patent: WO 01/2827-A 1 04-OCT-2001;  
 Millennium Pharmaceuticals, Inc. (US)  
 FEATURES  
 source Location/Qualifiers  
 1..2558  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 435..2321  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD10239.1"  
 /db\_xref="GI:16515795"  
 /translation="MAILPLLLCLPLAPASSPPPOSATPSPRCRCQTSLSPLSVL  
 CPAGLFLVPPSDDRAAELRLADNFIAVRRRLDLMNTGLLHLILSRNTHRHVAAGA  
 FADRLRALHLDGNRLTSLGECQLGLNHLIILSNOLAALAAAGLDLDCAEETLED  
 LDLSYNNLEQLPWEALRGUNVNTGLDHNHLLASVPAGAFSLHLKLARDLMTSNRLTT  
 IPDPLFSLRLPLALPRGSPASALVLAFGNPLHNCSELVWLRHLAREDDLEACASPP  
 ALGGRYEWAQVEEFVCEPVVTHRPPLAVPAPRALRCRAVGDPERVRWSPQG  
 RILGNSRRAFDPNLTLELLTEPFGDGIPTCIAANAAGEATAAEVLTVPVPPFQLA  
 NSTSDPRDGDALTPPSAASAKVADTPPTDRGVQVTEHGAALVAVQWDPORP  
 IPGRMVOIYQNSADDLVYRMI PAESRFLITDLASGRVTDLCVLAIVEGSAETGLT  
 ATRPVGCARFSTELAPRCAPAPPLGTMIIALGGVIVASVLVFIIVLLMRVYKHG  
 QGPQKAKIPAPSVSCVINGALGPTPTPAPPAPALRAHTVVRLLDCEPWPQGH  
 PVGP"  
 BASE COUNT 422 a 956 c 743 g 437 t  
 ORIGIN  
 Query Match 20.1%; Score 633; DB 6; Length 2558;  
 Best Local Similarity 61.3%; Pred. No. 7.5e-100;  
 Matches 1078; Conservative 0; Mismatches 665; Indels 16; Gaps 3;  
 507 GTTTCGCTGTGCGAGCGCTGCCCAAGTACTGTGTGCGCAGAACTGTCTGAGTCACT 566  
 497 GTACGCCACACCCAGCCCATGTCCCGCGCTGCCGCTGCCAGACACAGTCTGCCCT 556  
 567 GGGGACCTGTGCGCTCAAGGGCTGTCTTGTACCCCTGTATATGACCGCGGAC 626  
 557 AAGCGTGTGTGCCAGGGGAGGCTCTGTGTGTGTCACCTCGTGGACCGCGGC 616  
 627 AGTGGAGCTGCGCTCGGGCGGCAACTTCATCATCAATCAATCAATCAATCAATCAAT 686  
 617 AGCGGAGCTGCGCTGCGAGACAACTTCATCGCTCCGTCGCGCGCGCGCGCGCG 676  
 687 CATGACGGGCTGGTGGACCTGACCTGTCCAGGACACCATCAGCCACATCCAGCCCTT 746  
 677 CATGACAGGCTGCTGCATCTGAGCTGTGCGGGAACACATCCGCGAGTGGCTGCCGG 736  
 747 TTCTTTCTGAGCTGTGAGAGCTCGGCTCCCTGTGATCTTTGACAGAACTGCGCTGCC 806  
 737 CGCTTCGCGAGCTGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796  
 807 CCTTGGGAGGACACCTCCGCGGCTGTGTCACCTGTGAGCACTTATCTGCAACAA 866  
 797 ACTGGCGAGGGCAGCTGCGCGGCTGTGTCACCTGTGAGCACTTATCTGCAACAA 856  
 867 CCAGCTGGGCGGATCCAGATGAGGCTTTTGGAGACTTCTGCTGACATTTGAGGATCT 926  
 857 CCAGCTGGCGGCTGGCGGCGGCTGCTGATGATGATGATGATGATGATGATGATGATG 916  
 927 GGACCTCTCTTACAAACCTTCCATGGCTGCGGTGGGACTCGTGGGAGGATGCTCAA 986  
 917 CGACCTCTCTTACAAACCTTCCATGGCTGCGGTGGGAGGCTGCTGGGCGGCTGGG 976  
 987 CCTCCACAGCTGAGCTGAGCACCACTGCTGATCACTGCTGAGGAGGACCTTTGCTG 1046  
 977 CGTCAACAGCTGGGCGCTCGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036

1047 AGACCTGAGAAAATGGGCGCCCTGGATCTCAATCGGCTGCAGAGCTGCCCCC 1106  
 1037 CGCCCTGCACAGCTGGCGCCGCTGGACATACCTCCAAACCGCTGACCACAAATCCCA 1096  
 1107 TGATCCCATCTTTGCGCGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTTGGCCCA 1166  
 1097 CGACCACTCTTCTCCCGCTGCTGCCAGGCGCCGGGCTCGCGCGCTCTGCG 1156  
 1167 CTGTCTCTTTAGTTTGGGGTAACCCACTTCACTGCAATTTGAGCTTCTCTGCTGCG 1226  
 1157 CTGTGCTGCGCTTTTGGCGGAACCCCTGCTGCACTGCACTGCGAGCTGGTGTGCTG 1216  
 1227 GAGGCTCGAGCGGAGCATGACCTGGAACTGTGGCTTCCCGAGGGGCTCAAGGCTG 1286  
 1217 TCGCTTGGCGGGAGAGAGCTGAGGCTTGGCTGCGAGCGGCTCTCATCACCAGCA 1276  
 1287 CTACTTCTGGCATGTGCTGAGAGAGTTTGTGCGAGCGGCTCTCATCACCAGCA 1346  
 1277 CTACTTCTGGCGGTGGCGAGAGAGTTTGTGCGAGCGGCTCTCGCTGCGCGGCTG 1336  
 1347 CACACACAGTTCTGCTGTTCTGAGGCGGCGGCGGCGGCTCTCAAGTCAAGGCTTGG 1406  
 1337 CTACCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1396  
 1407 GGACCCAGCGCCCTTATCACTGGGTAGCCCGGATGACCGCTGCTGAGGAGCTCTTC 1466  
 1397 GGACCCAGCGCCCTTATCACTGGGTAGCCCGGATGACCGCTGCTGAGGAGCTCTTC 1456  
 1467 AAGGACCGTGTCTATGACAAATGACCGCTGAGCATCTTTCATCACCACATCTCAG 1526  
 1457 CCGTGGCGCGGCTTCCCAATGGGACGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 1516  
 1527 TGGTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1586  
 1517 TGGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1576  
 1587 TCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1644  
 1577 GACTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1636  
 1645 TCCCGCTCTCAGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704  
 1637 GACCGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1696  
 1705 GCGGAGAGCTTCCAAAGCCCCCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1764  
 1697 CGACACTGGGCGCCCTA-----CCGACCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1748  
 1765 ACCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1824  
 1749 ACAGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1808  
 1825 CTGCACTAACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1884  
 1809 ATCCAGTACAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1868  
 1885 AAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1944  
 1869 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1928  
 1945 ATGTGGGATGACACAGCAGACACTCAGGCGCACCAATCGTGGGCTGCGCCAGTTTC 2004  
 1929 GTGTATGAGGGAGTGCACGGGCTCAGGCGCACCGGCTGCTGCTGCTGCTGCTGCTGCT 1988  
 2005 TTACCAAGGCTGACTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2064  
 1989 TCCACCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2048  
 2065 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2124  
 2049 ATGATCATCGGCTGGGCGGCTCATCGTAGCCTCGGTACTGCTGCTGCTGCTGCTGCTG 2108

QY 2125 ATGGTCGCTACAGTCTGCAACACAGAGGGCCCCAG-----CAAGATGCGAGGGGCC 2178  
 Db 2109 CTAATCGCTACAAAGTGCACGGCGCCAGCCCGGCAAGGCTTCCCGCGGCC 2168  
 QY 2179 GTGAGCAATGTGTACTCGCAGACCAACGCGGCCCGCCCTCCCAAGCAGCGGCA 2238  
 Db 2169 GTTAGCAGGTTGTCTCCAGACCAACGCGGCCCTGGGCCCCACAGCCAGCGCGCCCG 2228  
 QY 2239 GCCGGGGCCCCCGCGAGG 2257  
 Db 2229 CCGCGCCCGAGCGCGCGG 2247

RESULT 14

BC015581 2458 bp mRNA linear PRI 17-JAN-2002  
 LOCUS Homo sapiens, clone MGC:23278 IMAGE:4637053, mRNA, complete cds.  
 ACCESSION BC015581  
 VERSION BC015581.1 GI:18182830  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 2458)  
 Strausberg,R.  
 Direct Submission  
 Submitted (04-OCT-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
 Tissue Procurement: DCTD/DTP/Gazdar  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
 Shevchenko,Y., Wechetby,K.D., Beckstrom-Sternberg,S.M.,  
 Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,  
 Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,  
 Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C.,  
 McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,  
 Tongson,E., Touchman,J.W., Tsurgan,C., Vogt,J.L., Walker,M.A.,  
 Zhang,L.-H. and Green,E.D.

REMARK COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 33 Row: C Column: 15  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Location/Qualifiers  
 1..2458  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:23278 IMAGE:4637053"  
 /tissue\_type="Lung, large cell carcinoma"  
 /clone\_lib="NIH MGC 18"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 1186..2160  
 /codon\_start=1  
 /product="Unknown (protein for MGC:23278)"  
 /protein\_id="AAH15581.1"  
 /db\_xref="GI:18182831"  
 /translation="MHVGPDRLVNCSRARAFNPTLIGVTGAGDAGGYTCIATN  
 PAGEATARVELRVLPALPHGNSAEGRGPSDIAASRTAAAGETLLESEPAVQTE  
 VTATSLGVSNGPGRPADPVMFQIQYNSEDETLIYRIVPASSHHFLKLHLPVPGADY

CDS

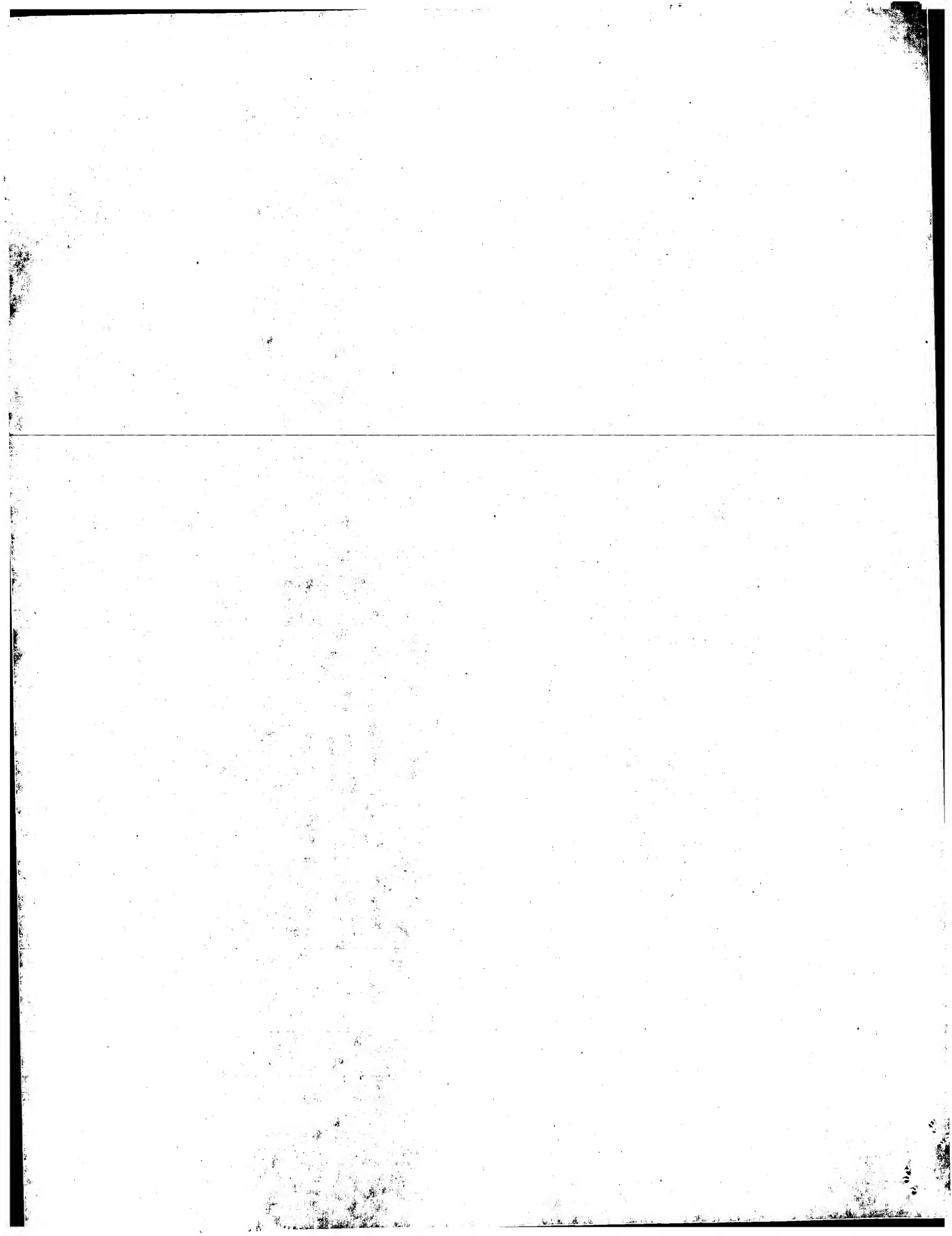
BASE COUNT 358 a 885 c 783 g 432 t  
 ORIGIN  
 Query Match 19.8%; Score 621; DB 9; Length 2458;  
 Best Local Similarity 62.2%; Pred. No. 9e-98;  
 Matches 1088; Conservative 0; Mismatches 610; Indels 51; Gaps 5;  
 QY 520 GACGCTGCCCCCAAGTACTGTGTGCGCAGAACTCTGTCTGAGTCACTTGGGGACCTGTGCG 579  
 Db 295 GCCGCTGCGCGCTGCTGCTGCGTCTGCAGAACTCTGCGAGTCTCGAGTCTCGACACCTCTGT 354  
 QY 580 CCTTCAAGGGGCTGCTCTTTGTATACCCCTTGATATTGACCGCGCAGCAGTGGAGCTGCGC 639  
 Db 355 GCCCACCAGGCTGCTGCTTTGTGCGGCCCAACGTTGGAGCGCGCAGCAGTGGAGCTGCGG 414  
 QY 640 CTGGCGGCAACTTTCATCATCATCATCAGCGCGCAGGACTTTGCCAACATGACGGGGCTG 699  
 Db 415 CTGGCTGACAACTTTCATCAGGCTTGGGGCGCCCTGACTTCCGCAACATGACGGGACTG 474  
 QY 700 GTGGACCTGACCTCTCCAGGAACACCATCAGCCACATCCAGCCCTTTTTCCTTTCTTGGAC 759  
 Db 475 GTGGACCTGACACTGTCTGCAATGCCATCACCAGCATTGGGGCGCGCCCTTTGGGAC 534  
 QY 760 CTCGAGACCTCCGCTCCCTGCTCATCTTGACAGCAATCGGCTGCCAAGCCTTTGGGAGGAC 819  
 Db 535 CTCGAGACCTGCGTGTCTCCACCTTGACGCAACAGCGTGGTGGAGCTGGGACCCGGG 594  
 QY 820 ACCCTCCGGGGCTGCTCAACTGCGACACCTTATCTGTGAACAACAACAGCAGCTGGGCGG 879  
 Db 595 AGCTTCCGGGGCGCGCTCAATCTGCGACACCTCATCTCAGCGGCAACAGCAGCTGGGCGG 654  
 QY 880 ATCGCAGATGAGGCTTTGAGGACTTCTGCTGACATTTGGAGGACTTGACCTCTCTTAC 939  
 Db 655 ATCGCGCGGAGGCTTTCGACACTTCTTAGAGAGCTGGAGGAGCTTGACCTGTCTTAC 714  
 QY 940 AACAACTCCATGGCTGCGCTGGGACTCCGTCGACGACATGGTCAACCTCCACACGCTG 999  
 Db 715 AACAACTCCGCGAGTGCCTTGGCGCGCATCGCGCCATGCTGCTGCTGACACCTC 774  
 QY 1000 AGCTTGGACCAACACTGCTGTGATCACATGCGCGAGGACACTTTGGAGACCTTGAGAAA 1059  
 Db 775 AACCTGGACCAATAACTTATTGACCACTTGCCCGGCGCGCTTCCGCGAGCTCGGTGAC 834  
 QY 1060 CTGGCGCGCTGATCTCACCTCCATCGGCTGCGAGAGCTGCCCGCTGATCCCATCTTT 1119  
 Db 835 CTCTCCGCGCTGAGCTCACCTCCAAACCGCTGGCGCAGCTGGCTCCGAGACCCGCTTTTC 894  
 QY 1120 GCCCGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTTGGCCCACTTTGTCTTTAGT 1179  
 Db 895 TCTGTTGGGCTGAT-----GCAGAGGCTCTCCCGCGCCCTTGTGCTGAGC 942  
 QY 1180 TTTGGGGTAACCCACTTCACTGCAATTTGTGAGTCTTCTTGGCTGGGAGGCTCGAGCGG 1239  
 Db 943 TTTAGCGGGAACCCCGCTGCACTGCACTGCTGAGCTGCTGCTGCTGCGGCGGCTGGCGCG 1002  
 QY 1240 GACGATGACCTGGAAACCTGTGGCTCCCGAGGGGCGCTCAAGGTCGCTACTTCTTGGCAT 1299  
 Db 1003 CCGGACGACCTTGGAAACGTGGCGCTTCCCGCGCGCGCTTGTGCTGCTTCTTGGGCA 1062  
 QY 1300 GTGCGTGGAGGAGGAGTTTGTGTGCGAGCGCGCTCTCATCACCCAGCACACACAAGTTG 1359  
 Db 1063 GTGCCCGAGGCGAGTTCTCTGTGAGCGCGCCCTCATTTGCCGCGCACACGAGGCGCTC 1122  
 QY 1360 CTGTTCTGGAGGCGGCGGCGCACACTCAAGTGCAGAAAGCATTTGGGAGCCCCAGCCGCC 1419  
 Db 1123 TGGTGTCTGGAAGGCGCAGCGGCGCACGCTGCGTGGCGCGCCCTGGGTGACCCCGCGCT 1182  
 QY 1420 CTTATCCACTGGGTAGCCCCCGATGACCGCTGGTAGGAACTCTCAAGGACCCCTGTC 1479  
 Db 1183 ACCATGCACTGGGTGGTCTCTGACGACCGGTTGGTTGGCAACTCTCTCCGAGCCCCGGCT 1242

UB-09-031-001-1:139

[illegible]

Search completed: June 8, 2003, 18:21:50  
Job time : 9339 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 15:09:17 ; Search time 664 Seconds  
(without alignments)  
10663.076 Million cell updates/sec

Title: US-09-831-846-1  
Perfect score: 3144  
Sequence: 1 gcctggctccctcctcgtga.....gcgggttgattgcttatg 3144

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125995159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3144	100.0	3144	21	AAA40083
2	3144	100.0	3144	21	AAA60605
3	2780.2	88.4	2818	22	AA544992
4	2780.2	88.4	2818	22	AA158215
5	969	30.8	969	23	AA577723
6	678.8	21.6	771	22	AA160001
7	634.6	20.2	1887	22	AAH25568
8	634.6	20.2	1887	22	AAH78215
9	633	20.1	1887	22	AA166986

10	633	20.1	2637	22	AA166985	Human leucine-rich
11	622.6	19.8	2369	24	AAAD38695	Human LP220 secret
12	621.8	19.8	2855	24	ABK70010	cDNA encoding huma
13	618.8	19.7	2316	22	AAH78204	Nucleotide sequenc
14	608	19.3	1653	24	AAAD28124	Human leucine-rich
15	554.6	17.6	1653	24	AAAD28125	Human leucine-rich
16	546	17.4	3012	24	ABN99416	Human secreted pro
17	544.4	17.3	2301	22	AAH78197	Nucleotide sequenc
18	477.2	15.2	1440	22	AAH78214	Nucleotide sequenc
19	477.2	15.2	8269	22	AAI99642	Human expressed po
20	477.2	15.2	8269	22	ABA06793	Human genomic DNA
21	477.2	15.2	8269	22	AAK90941	Human digestive sy
22	477.2	15.2	8269	22	AAI64010	Human polynucleoti
23	477.2	15.2	8269	22	AAAS31472	Human DNA for a no
24	477.2	15.2	8269	22	AAAS31976	Human liver associ
25	477.2	15.2	8269	22	AAAS33436	DNA encoding huma
26	477.2	15.2	8269	22	AAAS35069	DNA #19 encoding h
27	477.2	15.2	8269	24	ABQ66796	Human polynucleoti
28	477.2	15.2	8269	24	ABN90331	Human liver antige
29	471	15.0	1510	23	AAAS70133	DNA encoding novel
30	471	15.0	1602	22	AAH78203	Nucleotide sequenc
31	462.2	14.7	1582	23	AAAS69014	DNA encoding novel
32	461.4	14.7	1872	22	AAF76853	Human secreted pro
33	396	12.6	441	22	AAAS45180	cDNA encoding nove
34	298.2	9.5	1268	22	AAAS29522	Human endocrine po
35	226.2	7.2	541	22	AAI99624	Human expressed po
36	226.2	7.2	541	22	AAAS33219	DNA encoding huma
37	204.2	6.5	432	24	ABK33084	DNA encoding novel
38	204.2	6.5	432	24	AAAD28122	Human leucine-rich
39	194	6.2	902	22	ABA06401	Human cDNA SEQ ID
40	194	6.2	902	22	AAAS29645	Human endocrine po
41	194	6.2	902	22	AAI63902	Human polynucleoti
42	194	6.2	902	22	AAAS31239	Human cDNA encodin
43	194	6.2	902	22	AAAS34784	cDNA encoding nove
44	194	6.2	902	24	ABQ66563	Human polynucleoti
45	193.2	6.1	902	24	ABL90356	Human polynucleoti

ALIGNMENTS

RESULT 1  
AAA40083  
ID AAA40083 standard; cDNA; 3144 BP.  
XX  
XX  
AAA40083;  
AC  
XX  
DT 19-OCT-2000 (first entry)  
XX  
XX  
Human brain-specific transmembrane glycoprotein encoding cDNA.  
DE  
DE  
XX  
XX  
Transmembrane glycoprotein; human; brain; cellular signal transducer;  
KW neuroactive; neuroprotective; cerebroprotective; drug development;  
KW treatment; nervous disease; diagnostic; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 466..2835  
FT /\*tag= a  
FT /product= "transmembrane glycoprotein"  
XX  
XX  
WO200031256-A1.

XX  
PD 02-JUN-2000.  
XX  
XX  
PF 18-NOV-1999; 99WO-JP06449.  
XX  
XX  
PR 20-NOV-1998; 98JP-0331727.  
XX  
XX  
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
XX (KAZU-) KAZUSA DNA RES INST.



Db 1681 AGCCGGGAGGTGGAGGAGTGGGGCGGAGAGCCCTCCAAAAGCCCCCGGAAACGGGCT 1740  
Qy 1741 GTGCTTGTTGTTGAAGTACCAACCACTCGGCCCTGGTCAAGTGGTGTCTGACGAAGTCA 1800  
Db 1741 GTGCTTGTTGTTGAAGTACCAACCACTCGGCCCTGGTCAAGTGGTGTCTGACGAAGTCA 1800  
Qy 1801 GCACCCCGGTTGAAGTGTACAGTGCAGTGAACAAGTGTCTGACGATGAGTACTGATT 1860  
Db 1801 GCACCCCGGTTGAAGTGTACAGTGCAGTGAACAAGTGTCTGACGATGAGTACTGATT 1860  
Qy 1861 TACAGGATGATCCAGGCTCCAAAGGCTTTCGTGGTCAACAACCTGTGTGTCAGGACT 1920  
Db 1861 TACAGGATGATCCAGGCTCCAAAGGCTTTCGTGGTCAACAACCTGTGTGTCAGGACT 1920  
Qy 1921 GGTCTAGGATGTTGTGTGTCGTCATGTCGTCATGTCGTCATGTCGTCATGTCGTCATG 1980  
Db 1921 GGTCTAGGATGTTGTGTGTCGTCATGTCGTCATGTCGTCATGTCGTCATGTCGTCATG 1980  
Qy 1981 AACATGTTGGGTCGCGCCAGTTCACCAAGGCTGACTACCCGAGTGCAGTCCATG 2040  
Db 1981 AACATGTTGGGTCGCGCCAGTTCACCAAGGCTGACTACCCGAGTGCAGTCCATG 2040  
Qy 2041 CACAGCCAGATTCGCGCGGCGCACCATGATCTGTTGTCATCGGGGCGATCATCTGTCGCG 2100  
Db 2041 CACAGCCAGATTCGCGCGGCGCACCATGATCTGTTGTCATCGGGGCGATCATCTGTCGCG 2100  
Qy 2101 CTGCTGGTCTTCATCTCATCTCTCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2160  
Db 2101 CTGCTGGTCTTCATCTCATCTCTCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2160  
Qy 2161 ACAGAGATGGAGCGGCGGTGAGCAATGTGTACTGTCAGACCAACGGGCGCCAGCCACCG 2220  
Db 2161 ACAGAGATGGAGCGGCGGTGAGCAATGTGTACTGTCAGACCAACGGGCGCCAGCCACCG 2220  
Qy 2221 CTTCAAGCAGCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2280  
Db 2221 CTTCAAGCAGCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2280  
Qy 2281 AACAGAGCTCTGAGCTTACACCGCCAGCTGGCGCGGCGGCGGCGGCGGCGGCGGCGG 2340  
Db 2281 AACAGAGCTCTGAGCTTACACCGCCAGCTGGCGCGGCGGCGGCGGCGGCGGCGGCGG 2340  
Qy 2341 TCCCTGGGAGTGGGAGGCTCGGGGCTGGAGAGCGGGCGGCGGCGGCGGCGGCGGCGGCGG 2400  
Db 2341 TCCCTGGGAGTGGGAGGCTCGGGGCTGGAGAGCGGGCGGCGGCGGCGGCGGCGGCGGCGG 2400  
Qy 2401 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2460  
Db 2401 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2460  
Qy 2461 AAGAGTCAGAGAAAGGAGGAGTGTGTGACTCCAGGACTCCAGCGGAGAGGGGCTGGG 2520  
Db 2461 AAGAGTCAGAGAAAGGAGGAGTGTGTGACTCCAGGACTCCAGCGGAGAGGGGCTGGG 2520  
Qy 2521 AGTTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2580  
Db 2521 AGTTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2580  
Qy 2581 GCCAGAGCTGCTCCCTTCGCTTGGAGGCGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2640  
Db 2581 GCCAGAGCTGCTCCCTTCGCTTGGAGGCGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2640  
Qy 2641 ATGGGGGACTTTGCTGCTGCGCGGCGGAGGGGCTGTCGCGGGGGGCTACAGTCTCTCT 2700  
Db 2641 ATGGGGGACTTTGCTGCTGCGCGGCGGAGGGGCTGTCGCGGGGGGCTACAGTCTCTCT 2700  
Qy 2701 CGGAAGGTCTGAAACATCTGGAAGCGAGCGAGCTCTCTGTCGTCGTCGTCGTCGTCGTCG 2760  
Db 2701 CGGAAGGTCTGAAACATCTGGAAGCGAGCGAGCTCTCTGTCGTCGTCGTCGTCGTCGTCG 2760  
Qy 2761 TTTGAGGAGAGTACCTGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2820  
Db 2761 TTTGAGGAGAGTACCTGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2820

Qy 2821 GAGACACGGTCTAGGTGGGGTGGCATGCTCTTCTCTGTCGCGAGGTGGGAGAAG 2880  
Db 2821 GAGACACGGTCTAGGTGGGGTGGCATGCTCTTCTCTGTCGCGAGGTGGGAGAAG 2880  
Qy 2881 GGGAAAGATCTCACTGGCAAGTGTGTTGGAGTTTCCATGGTGTATGTTTACATCCAGG 2940  
Db 2881 GGGAAAGATCTCACTGGCAAGTGTGTTGGAGTTTCCATGGTGTATGTTTACATCCAGG 2940  
Qy 2941 ACAGTTTCTGCTCTCCCTGTCAATGGCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 3000  
Db 2941 ACAGTTTCTGCTCTCCCTGTCAATGGCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 3000  
Qy 3001 CCTCCCAACCCAGCGGCGGCTGCTCAGGGAATGTGGACTCGCTCAAAATGCCGACT 3060  
Db 3001 CCTCCCAACCCAGCGGCGGCTGCTCAGGGAATGTGGACTCGCTCAAAATGCCGACT 3060  
Qy 3061 GAGCCCTCAGTGTGTTGAAAGCGGAGACTCCGCTTCTTAATCAGAAATGTAGCTTACAA 3120  
Db 3061 GAGCCCTCAGTGTGTTGAAAGCGGAGACTCCGCTTCTTAATCAGAAATGTAGCTTACAA 3120  
Qy 3121 GCAAGCGGCTTTGGATTGCTTATG 3144  
Db 3121 GCAAGCGGCTTTGGATTGCTTATG 3144  
  
RESULT 2  
AAA60605  
ID AAA60605 standard; cDNA; 3144 BP.  
XX  
AC AAA60605;  
XX  
DT 19-OCT-2000 (first entry)  
XX  
DE Human hh00149 protein encoding cDNA SEQ ID NO:3.  
XX  
KW Human; ubiquitin-like protein; 149Y2H#151; hh00149; brain;  
KW 2-hybrid screening; neuroprotective; signal transducer;  
KW nervous system disease; diagnosis; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 466..2835  
FT /\*tag= a  
FT /product= "hh00149"  
XX  
PN WO200031255-A1.  
XX  
PD 02-JUN-2000.  
XX  
PF 18-NOV-1999; 99WO-JP06448.  
XX  
PR 20-NOV-1998; 98JP-0331701.  
XX  
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
XX  
PI Funahashi S, Miyata S;  
XX  
WP1; 2000-400066/34.  
DR P-PSDB; AAB12448.  
XX  
PT Gene encoding ubiquitin-like protein which interacts with protein  
PT hh00149, useful in the diagnosis and treatment of diseases associated  
PT with the nervous system -  
XX  
PS Example 2; Page 64-73; 88pp; Japanese.  
XX  
CC The present invention describes a ubiquitin-like protein, designated  
CC 149Y2H#151, which interacts with protein hh00149 expressed specifically  
CC in the brain. The 149Y2H#151 protein has neuroprotective activity, and  
CC is a signal transducer. The 149Y2H#151 gene and encoded protein are  
CC useful in the diagnosis and treatment of diseases associated with the



Db 3121 GCAAGCGCTTGGATTGCTTATG 3144

|||||

RESULT 3

AAS4992

XX AAS4992 standard; cDNA; 2818 BP.

XX AC AAS4992;

XX DT 18-DEC-2001 (first entry)

XX DE cDNA encoding novel human secretory protein, Seq ID No 73.

XX KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock; ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy; transgenic animal; Alzheimer's disease; Parkinson's disease; burn; amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia; ulcer; osteoporosis; bone degenerative disorder; periodontal disease; gut protection; lung; liver fibrosis; immune deficiency; infection; severe combined immunodeficiency; SCID; autoimmune disorder; allergy; multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma; fertility; analgesic; pain; antigen; ss.

XX OS Homo sapiens.

XX PN WO200166689-A2.

XX PD 13-SEP-2001.

XX PF 05-MAR-2001; 2001WO-US04942.

XX PR 07-MAR-2000; 2000US-0519705.

XX PR 19-MAY-2000; 2000US-0574454.

XX PR 17-JUN-2000; 2000US-0596193.

XX PR 14-JUL-2000; 2000US-0616847.

XX PR 19-SEP-2000; 2000US-0665363.

XX PR 20-OCT-2000; 2000US-0693267.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

XX PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX WPI; 2001-589934/66.

XX P-PSDB; AAU28092.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis and treatment of cancer, neurological, inflammatory, and autoimmune disorders -

XX Claim 1; SEQ ID No 73; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (I) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (I) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (I) is also useful for

gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAS44920-AAS45295 represent novel human secreted protein coding sequences of the invention.

X C  
sequence 2818 BP: 523 A: 933 C: 825 G: 537 T: 0 other;

Query Match 88.4%: Score 2780.2; DB 22; Length 2818;

Query Match	80.4%	Sec. No. 0;
Best Local Similarity	99.7%	Pred. NO. 0;
Matches 2785; Conservative		8; Indels
		0; Mismatches
		0; Gaps
		0;

43	QY	CACAA	CGCGG	CAGGCT	CGTCT	GAACTT	GAA	CACAC	CCCC	CACAT	TCAAG	ATG	CCCGAG	TT	102
26	Db	CCCC	CCCC	CTGG	CTCGT	CTG	CTG	AACTT	GAA	CAC	CCCA	CACAT	TCAAG	ATG	85
103	QY	CTTGG	AAATG	CTCGGG	TTCTT	CGATC	CGG	AAAT	TCTA	CGG	CAT	CCTCT	CCTAGG	GAGG	162
86	Db	CCTGG	AAATG	CTCGGG	TTCTT	CGATC	CGG	AAAT	TCTA	CGG	CAT	CCTCT	CCTAGG	GAGG	145
163	QY	ATTAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	222
146	Db	ATTAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	205
223	QY	CTTCT	TTCTT	GTCT	AAAC	CCCAT	GAT	AATA	AACT	GAAT	TAAG	CGCT	CAGG	CTT	282
206	Db	CTTCT	TTCTT	GTCT	AAAT	TGCC	ATG	ATAT	AAT	TAACT	GAAT	TAAG	CGCT	CAGG	265
283	QY	GGAC	GTCC	GAGG	CCAC	ATCT	GTG	CTG	CAAT	TCC	CGG	AGCC	CGCGG	AGG	342
266	Db	GGAC	GTCC	GAGG	CCAC	ATCT	GTG	CTG	CAAT	TCC	CGG	AGCC	CGCGG	AGG	325
343	QY	TCTGT	CTCGG	CGGGA	AGG	ATCG	TGG	CCG	AGCC	CGG	GAGCC	CGG	CGCC	CGCGG	402
326	Db	TCTGT	CTCGG	CGGGA	AGG	ATCG	TGG	CCG	AGCC	CGG	GAGCC	CGG	CGCC	CGCGG	385
403	QY	CGG	CCCTCG	GTC	CCAC	CAGC	CGG	GGGT	AGAT	GTG	CTG	CCG	AGCG	CTG	462
386	Db	CGG	CCCTCG	GTC	CCAC	CAGC	CGG	GGGT	AGAT	GTG	CTG	CCG	AGCG	CTG	445
463	QY	ACCAT	GAG	AGC	CCCTG	CTT	GTG	GCCT	GTCT	AG	CGTT	TGG	CAT	CGCT	522
446	Db	ACCAT	GAG	AGC	CCCTG	CTT	GTG	GCCT	GTCT	AG	CGTT	TGG	CAT	CGCT	505
523	QY	GCCT	GTCCC	CAAG	TACT	GTCT	GTG	CA	GAAT	CTGT	CTG	AGT	CACT	TGGG	582
506	Db	GCCT	GTCCC	CAAG	TACT	GTCT	GTG	CA	GAAT	CTGT	CTG	AGT	CACT	TGGG	565
583	QY	TCCA	AGGG	GTG	CTCT	TTG	TAC	CCCT	GTAT	TGAC	CGG	GGAC	AGT	TGAG	642
566	Db	TCCA	AGGG	GTG	CTCT	TTG	TAC	CCCT	GTAT	TGAC	CGG	GGAC	AGT	TGAG	625
643	QY	GGCG	CAACT	TTCAT	CCAT	CCAC	TAC	CAG	CGCC	AGG	ACTTT	TGCC	AAAT	GACGGG	702
626	Db	GGCG	CAACT	TTCAT	CCAT	CCAC	TAC	CAG	CGCC	AGG	ACTTT	TGCC	AAAT	GACGGG	685
703	QY	GAC	CTG	ACCT	TGT	CCAG	GAA	CA	CCAT	CAG	CCAT	CCAG	CCCTTT	CTTT	762
686	Db	GAC	CTG	ACCT	TGT	CCAG	GAA	CA	CCAT	CAG	CCAT	CCAG	CCCTTT	CTTT	745
763	QY	GAG	AGC	CTCC	CGCT	CCCT	GCAT	CTTT	GAC	AGAA	TGG	CGT	CCAG	CCCT	822
746	Db	GAG	AGC	CTCC	CGCT	CCCT	GCAT	CTTT	GAC	AGAA	TGG	CGT	CCAG	CCCT	805
823	QY	CTTC	CGG	CGC	CTT	GTCT	CACT	GTG	TGAA	CAAC	CAAC	CAG	CTTGG	CGG	882



Db 1886 AACCTGGTGTACAGGACTGGCTACGACTTGTGTGTCTGGCCATGTGGGATGACACAGCC 1945  
Qy 1963 ACGACACTCAGCCACCAACATCTGTGGCTGCGCCAGTCTTTCACCAAGGCTGACTAC 2022  
Db 1946 ACGACACTCAGCCACCAACATCTGTGGCTGCGCCAGTCTTTCACCAAGGCTGACTAC 2005  
Qy 2023 CCGAGTGCAGTCTCATGACAGCCAGATCTTGGGCGGCACCATGATCTGTGTCATCGGG 2082  
Db 2006 CCGAGTGCAGTCTCATGACAGCCAGATCTTGGGCGGCACCATGATCTGTGTCATCGGG 2065  
Qy 2083 GGCATCATCTGGCCAGCTGTGTCTTCATCTCATCTCTCATCTGTGCTGCTACAGGTC 2142  
Db 2066 GGCATCATCTGGCCAGCTGTGTCTTCATCTCATCTCTCATCTGTGCTGCTACAGGTC 2125  
Qy 2143 TGCAACACAGGCCCCCAGCAGATGGCAGGCGCGTGGCAATGTGTACTCGCAGACC 2202  
Db 2126 TGCAACACAGGCCCCCAGCAGATGGCAGGCGCGTGGCAATGTGTACTCGCAGACC 2185  
Qy 2203 AACGGGCCCCAGCCAGCTCTCAAGCAGGACACAGCGGGCCCCCGCAGGCGCCG 2262  
Db 2186 AACGGGCCCCAGCCAGCTCTCAAGCAGGACACAGCGGGCCCCCGCAGGCGCCG 2245  
Qy 2263 CCGAAGGTGTGTGGCGCAACGAGCTCTTGGACTTCACCGCCAGCTGCGCGCCGCACT 2322  
Db 2246 CCGAAGGTGTGTGGCGCAACGAGCTCTTGGACTTCACCGCCAGCTGCGCGCCGCACT 2305  
Qy 2323 GACTCTCTTCTCCAGCTCCCTGGCAGTGGGAGGCTGCGGGCTGGGAGCGGGCCCC 2382  
Db 2306 GACTCTCTTCTCCAGCTCCCTGGCAGTGGGAGGCTGCGGGCTGGGAGCGGGCCCC 2365  
Qy 2383 TGGAGATCCACCTTCCCGCCCGCCCAAGCCAGCCTTGACCCCTGTATGGGGCC 2442  
Db 2366 TGGAGATCCACCTTCCCGCCCGCCCAAGCCAGCCTTGACCCCTGTATGGGGCC 2425  
Qy 2443 TTCGCTCTTCTCCAGCTCAGAGTCAAGAGGAGGAGTCTGAGCTCCAGGACTCA 2502  
Db 2426 TTCGCTCTTCTCCAGCTCAGAGTCAAGAGGAGGAGTCTGAGCTCCAGGACTCA 2485  
Qy 2503 GCGGGAGAGGGCTGGGAGCTGCGCCCGGGCCACCACTCGGACCGAGAGCACTGTCTG 2562  
Db 2486 GCGGGAGAGGGCTGGGAGCTGCGCCCGGGCCACCACTCGGACCGAGAGCACTGTCTG 2545  
Qy 2563 GGGCCCCCTTGGGCCCCGAGGAGCTGCTCCCTTCCCTTGGAGGGCAAGGCCAAA 2622  
Db 2546 GGGCCCCCTTGGGCCCCGAGGAGCTGCTCCCTTCCCTTGGAGGGCAAGGCCAAA 2605  
Qy 2623 CGCAGCCACTCTTGCATGGGAGCTTGTGCTGCGCGCGGAGGGCTGTCGCG 2682  
Db 2606 CGCAGCCACTCTTGCATGGGAGCTTGTGCTGCGCGCGGAGGGCTGTCGCG 2665  
Qy 2683 GCGGCTACAGTCTCTCGGAAGGTCTCGAATCTGGACGAGGCGGAGCTCTCTGTC 2742  
Db 2666 GCGGCTACAGTCTCTCGGAAGGTCTCGAATCTGGACGAGGCGGAGCTCTCTGTC 2725  
Qy 2743 AACGGCATCTTGTCCCTTTGAGGAGAGTACCTGGTGGGGCCCCGGGGGACTTTTGGC 2802  
Db 2726 AACGGCATCTTGTCCCTTTGAGGAGAGTACCTGGTGGGGCCCCGGGGGACTTTTGGC 2785  
Qy 2803 AGCTCCGATGGGTGATGAGAGCAGCGTCTAG 2835  
Db 2786 AGCTCCGATGGGTGATGAGAGCAGCGTCTAG 2818

RESULT 4  
AA158215  
ID AA158215 standard; cDNA; 2818 BP.  
AC  
AA158215;  
XX  
22-OCT-2001 (first entry)  
DT  
XX  
DE Human polynucleotide SEQ ID NO 418.  
XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.  
XX Homo sapiens.  
XX WO200153312-A1.  
PD 26-JUL-2001.  
XX 26-DEC-2000; 2000WO-US34263.  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J; Zhao QA, Zhou P, Goodrich R, Dmanac RT; WPI; 2001-442253/47.  
DR P-PSDB; AAM39059.  
XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -  
PT Claim 1; SEQ ID NO 418; 10078pp; English.  
XX The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukaemias and CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed specification.  
XX  
SQ Sequence 2818 BP; 523 A; 933 C; 825 G; 537 T; 0 other;

Query Match 88.4%; Score 2780.2; DB 22; Length 2818;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2785; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 43 CACAACCCCGGAGGCTCGTCTGAACCTTGAAGACACCCACACATTCACAGATGCCGAGGTT 102  
Db 26 CCCCCCCCCCTGGCTCGTCTGAACCTTGAAGACACCCACATTCACAGATGCCGAGGTT 85'  
Qy 103 CTTGGGAATGCTGGGGTTCTTCGATCCGGAATAATCTTACCGGATCTCTCTAGGAGGG 162  
Db 86 CTTGGGAATGCTGGGGTTCTTCGATCCGGAATAATCTTACCGGATCTCTCTAGGAGGG 145  
Qy 163 ATTAT 222  
Db 146 ATTAT 205

QY 223 CTTCTTCTGCTAAAGCCCATCGATATAACTGAATAAGGGCTCAGGCTTTCCCGCGT 282  
DB 206 CTTCTTCTTGTAAATGCGCATGATATAACTGAATAAGGGCTCAGGCTTTCCCGCGT 265  
QY 283 GGACGTCGAGGCCACCACTCTGCTGCAATTCGCCGGAGCGCGGAGGGTTTAGCTCGAG 342  
DB 266 GGACGTCGAGGCCACCACTCTGCTGCAATTCGCCGGAGCGCGGAGGGTTTAGCTCGAG 325  
QY 343 TCTGTCGCGGGGGAAGGATGCGTGGCGAGCGGGAGCGCGGGGCGCGCGGGAGC 402  
DB 326 TCTGTCGCGGGGGAAGGATGCGTGGCGAGCGGGAGCGCGGGGCGCGCGGGAGC 385  
QY 403 CGGCTCGTGCACCCAGCGGGGTAGATGCTGCTCGCCAGCGCGGTGAGTCACACAG 462  
DB 386 CGGCTCGTGCACCCAGCGGGGTAGATGCTGCTCGCCAGCGCGGTGAGTCACACAG 445  
QY 463 ACCATGGAGACCCCTGCTTGGTGGCTGCTAGGCTTTGGCATGCGCTTTGCCGTGTCGAC 522  
DB 446 ACCATGGAGACCCCTGCTTGGTGGCTGCTAGGCTTTGGCATGCGCTTTGCCGTGTCGAC 505  
QY 523 GCCTGCCCAAGTACTGTCGCCAGAACTCTGTCTGAGTCACTGGGACCCCTGTGCCCTG 582  
DB 506 GCCTGCCCAAGTACTGTCGCCAGAACTCTGTCTGAGTCACTGGGACCCCTGTGCCCTG 565  
QY 583 TCCAGGGGCTGCTTTGTACCCCTGATATTGACCGCGGACAGTGGAGCTGGCCCTG 642  
DB 566 TCCAGGGGCTGCTTTGTATCCCTGATATTGACCGCGGACAGTGGAGCTGGCCCTG 625  
QY 643 GCGGCAACTTCATCATCCATCAGCGCGCCAGGACTTTGCCAACAATGACGGGGCTGGTG 702  
DB 626 GCGGCAACTTCATCATCCATCAGCGCGCCAGGACTTTGCCAACAATGACGGGGCTGGTG 685  
QY 703 GACTGACCTGTCAGGAAACACATCAGCCACATCGAGCCCTTTTCCTTCTCGACCTC 762  
DB 686 GACTGACCTGTCAGGAAACACATCAGCCACATCGAGCCCTTTTCCTTCTCGACCTC 745  
QY 763 GAGAGCTCCGCTCCCTGCTGCTTTGACGCAATCGGCTGCAAGCCCTGGGGAGGACAC 822  
DB 746 GAGAGCTCCGCTCCCTGCTGCTTTGACGCAATCGGCTGCAAGCCCTGGGGAGGACAC 805  
QY 823 CTCGGGGCTGGTCAAACCTGACGACCTTTATCGTGAACAACAACAGCTGGCGGCATC 882  
DB 806 CTCGGGGCTGGTCAAACCTGACGACCTTTATCGTGAACAACAACAGCTGGCGGCATC 865  
QY 883 GCAGATGAGCTTTGAGGACTTCTGCTGACATTTGAGGATCTGGACCTCTCTCAAC 942  
DB 866 GCAGATGAGCTTTGAGGACTTCTGCTGACATTTGAGGATCTGGACCTCTCTCAAC 925  
QY 943 AACCTCCATGGCTCGGCTGGGACTCGCTGCGAGCATGGTCAACCTCCACAGCTGAGC 1002  
DB 926 AACCTCCATGGCTCGGCTGGGACTCGCTGCGAGCATGGTCAACCTCCACAGCTGAGC 985  
QY 1003 CTGGACCAACAACCTGCTGGATCAATCGCGAGGGCACTTTGAGACCTGCGAAGACTG 1062  
DB 986 CTGGACCAACAACCTGCTGGATCAATCGCGAGGGCACTTTGAGACCTGCGAAGACTG 1045  
QY 1063 GCCCGCTGATCTCACTCCATCGCTGAGAGCTGCCCTGATCCATCTTTGCC 1122  
DB 1046 GCCCGCTGATCTCACTCCATCGCTGAGAGCTGCCCTGATCCATCTTTGCC 1105  
QY 1123 CGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTGCCCGACCCCTTTGCTTTAGTTT 1182  
DB 1106 CGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTGCCCGACCCCTTTGCTTTAGTTT 1165  
QY 1183 GGGGTAACCCACTTCACTGCAATTTGAGTCTCTGCTGCGAGGCTCGAGCGGAC 1242  
DB 1166 GGGGTAACCCACTTCACTGCAATTTGAGTCTCTGCTGCGAGGCTCGAGCGGAC 1225  
QY 1243 GATGACCTGGAACCTGCTGGCTCCCGAGGGGCTCAAGGCTCGCTACTTCTGGCATGTG 1302  
DB 1226 GATGACCTGGAACCTGCTGGCTCCCGAGGGGCTCAAGGCTCGCTACTTCTGGCATGTG 1285  
QY 1303 CGTGAGGAGGATTTGTGTGCGAGCGGCTCTCATACCCAGCACACACAAGTTGCTG 1362

DB 1286 CGTGAGGAGGATTTGTGCGAGCGGCTCTCATCCCGACACACACAAGTTGCTG 1345  
QY 1363 GTTCTGAGGGCCAGGGCCACACTCAAGTGAAAGCCATTGGGAGACCCAGCCCCCTT 1422  
DB 1346 GTTCTGAGGGCCAGGGCCACACTCAAGTGAAAGCCATTGGGAGACCCAGCCCCCTT 1405  
QY 1423 ATCCACTGGGTAGCCCCCGATGACCGCTGCTGTAAGGAACTCTCAAGGACCGTGTCTAT 1482  
DB 1406 ATCCACTGGGTAGCCCCCGATGACCGCTGCTGTAAGGAACTCTCAAGGACCGTGTCTAT 1465  
QY 1483 GACAAATGGCACCCTGAGACATCTTCAATCACACATCTCAGACAGTGTGCTTCACTGC 1542  
DB 1466 GACAAATGGCACCCTGAGACATCTTCAATCACACATCTCAGACAGTGTGCTTCACTGC 1525  
QY 1543 ATTCTGCCAATGCTGCGGAGAGGCCACCGCCATGCTGAGGCTCTCAATCGTCCAGCTG 1602  
DB 1526 ATTCTGCCAATGCTGCGGAGAGGCCACCGCCATGCTGAGGCTCTCAATCGTCCAGCTG 1585  
QY 1603 CCACACCTCAGCAACAGCAGCCGCACTGCACCCCCCAAGTCCCGCTCTCAGACATC 1662  
DB 1586 CCACACCTCAGCAACAGCAGCCGCACTGCACCCCCCAAGTCCCGCTCTCAGACATC 1645  
QY 1663 ACTGCTCCAGCAACAGCAGCCGGGAGGTGGAGCAGTGGGGGGAGAGCTCCCAA 1722  
DB 1646 ACTGCTCCAGCAACAGCAGCCGGGAGGTGGAGCAGTGGGGGGAGAGCTCCCAA 1705  
QY 1723 AGCCCCCGGAACGGGCTGCTTGTCTGAACTGACCAACCACTCGGGCCCTGGTCAAG 1782  
DB 1706 AGCCCCCGGAACGGGCTGCTTGTCTGAACTGACCAACCACTCGGGCCCTGGTCAAG 1765  
QY 1783 TGGTCTGTGCAAGTCAAGCAACCCCGGTGAAGATGTACCAAGTGCAGTCAAACTGCTCT 1842  
DB 1766 TGGTCTGTGCAAGTCAAGCAACCCCGGTGAAGATGTACCAAGTGCAGTCAAACTGCTCT 1825  
QY 1843 GACGATGAGTACTGATTTACAGGATGATCCAGCTCCCAAGGCTTCGTGGTCAAC 1902  
DB 1826 GACGATGAGTACTGATTTACAGGATGATCCAGCTCCCAAGGCTTCGTGGTCAAC 1885  
QY 1903 AACCTGGTGTCAAGGACTGGCTTACGACTTGTGTGTGTGGCCATGTGGATGACACGCC 1962  
DB 1886 AACCTGGTGTCAAGGACTGGCTTACGACTTGTGTGTGTGGCCATGTGGATGACACGCC 1945  
QY 1963 ACACACTCAAGGCCCAACCAATCGTGGCTGGCCAGTCTTCCACCAAGGCTGACTAC 2022  
DB 1946 ACACACTCAAGGCCCAACCAATCGTGGCTGGCCAGTCTTCCACCAAGGCTGACTAC 2005  
QY 2023 CGGAGTGCAGTCCATGCAGAGCAGATTTGGGCGGCACCATGATCTGTGCTATCGGG 2082  
DB 2006 CGGAGTGCAGTCCATGCAGAGCAGATTTGGGCGGCAGATTTGGGCGGCACCATGATCTGTGCTATCGGG 2065  
QY 2083 GGCATCATCGTGGCCACGCTGCTGCTTTCATCGTCACTCTCATGGTGCCTTACAAGTTC 2142  
DB 2066 GGCATCATCGTGGCCACGCTGCTGCTTTCATCGTCACTCTCATGGTGCCTTACAAGTTC 2125  
QY 2143 TGCAACCAAGGAGGCCCAAGCAATGGCAGCGGCTGAGCAATGTGTACTCGAGACC 2202  
DB 2126 TGCAACCAAGGAGGCCCAAGCAATGGCAGCGGCTGAGCAATGTGTACTCGAGACC 2185  
QY 2203 AACGGGCGCCAGCCACCGCTCCAAAGCAGCAGCAGCGGGGCGCCCGCGAGGGCGCG 2262  
DB 2186 AACGGGCGCCAGCCACCGCTCCAAAGCAGCAGCAGCGGGGCGCCCGCGAGGGCGCG 2245  
QY 2263 CGGAAGGTGGTGGCGCAACAGCTCTTGGATTTCACCGCAGGCTTGGCCCGCGGCGAGT 2322  
DB 2246 CGGAAGGTGGTGGCGCAACAGCTCTTGGATTTCACCGCAGGCTTGGCCCGCGGCGAGT 2305  
QY 2323 GACTCTCTTCTTCCAGCTCCCTGGAGTGGGGGCTGCGGGGTGGAGGGGCCCC 2382  
DB 2306 GACTCTCTTCTTCCAGCTCCCTGGAGTGGGGGCTGCGGGGTGGAGGGGCCCC 2365  
QY 2383 TGGAGGATCCACCTTCGCCCGCGCCCAAGCCAGGCTTTGACCCCTGTATGGGGGC 2442

Db 2366 TGGAGATCCACCTCCCGCCGCGCCCAAGCCAGCCCTTGACCGCCTGATGGGGCC 2425  
Qy 2443 TTCGCTCCCTGACCTCAAGAGTCAGAGAAAGGAGAGCTGCTGGACTCCAGGACTCCA 2502  
Db 2426 TTCGCTCCCTGACCTCAAGAGTCAGAGAAAGGAGAGCTGCTGGACTCCAGGACTCCA 2485  
Qy 2503 GCGGGAGAGGGCTGGAGCTGCGCCCGGGGCGCCACTCGGACCGAGAGCCACTGCTG 2562  
Db 2486 GCGGGAGAGGGCTGGAGCTGCGCCCGGGGCGCCACTCGGACCGAGAGCCACTGCTG 2545  
Qy 2563 GGGCCCCCTGCGCGCGCGGCGAGGAGCTGCTCCCTTGGCGCTGGAGGCGAGGCCAAA 2622  
Db 2546 GGGCCCCCTGCGCGCGCGGCGAGGAGCTGCTCCCTTGGCGCTGGAGGCGAGGCCAAA 2605  
Qy 2623 CGCAGCACTCTTCGACATGGGGACTTTGCTGCGCGCGGCGAGGGTCTGTCGCG 2682  
Db 2606 CGCAGCACTCTTCGACATGGGGACTTTGCTGCTGCGCGCGGCGAGGGTCTGTCGCG 2665  
Qy 2683 GCGCGCTACAGTCTCTCGGAAGGTCTCGAACATCTGGACGAAGCGAGCCTCTCTGTC 2742  
Db 2666 GCGCGCTACAGTCTCTCGGAAGGTCTCGAACATCTGGACGAAGCGAGCCTCTCTGTC 2725  
Qy 2743 AACGCACTCTTTCGCTTTGAGGAGAGTGACCTGGTGGGCGCGCGGGGACTTTTGGC 2802  
Db 2726 AACGCACTCTTTCGCTTTGAGGAGAGTGACCTGGTGGGCGCGCGGGGACTTTTGGC 2785  
Qy 2803 AGCTCCGAATGGGTGATGAGAGCACGGTCTAG 2835  
Db 2786 AGCTCCGAATGGGTGATGAGAGCACGGTCTAG 2818  
RESULT 5  
ID AAS77723 standard; cDNA; 969 BP.  
XX AAS77723;  
AC AAS77723;  
XX 13-FEB-2002 (first entry)  
DT DNA encoding novel human diagnostic protein #13527.  
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX Homo sapiens.  
XX WO200175067-A2.  
XX 11-OCT-2001.  
XX 30-MAR-2001; 2001WO-US08631.  
XX 31-MAR-2000; 2000US-0540217.  
XX 23-AUG-2000; 2000US-0649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
XX P-PSDB; ABG13536.  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 1; SEQ ID No 13527; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 969 BP; 166 A; 330 C; 316 G; 157 T; 0 other;  
Query Match 30.8%; Score 969; DB 23; Length 969;  
Best Local Similarity 100.0%; Pred. No. 1.4e-194;  
Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1867 ATGATCCAGGCTCCAAAGAGCTTCGTGGTCAACACCTTCGTGGTCAAGGACTGGCTAC 1926  
Db 1 ATGATCCAGGCTCCAAAGAGCTTCGTGGTCAACACCTTCGTGGTCAAGGACTGGCTAC 60  
Qy 1927 GACTTGTGTGCTGGCCATGTGGATGACACAGCAGCAGCACTCACGCCACCAACATC 1986  
Db 61 GACTTGTGTGCTGGCCATGTGGATGACACAGCAGCAGCACTCACGCCACCAACATC 120  
Qy 1987 GTGGGCTGGCCAGCTTCTTACCAAGGCTGACTACCGCGAGTGCAGTCCATGCACAGC 2046  
Db 121 GTGGGCTGGCCAGCTTCTTACCAAGGCTGACTACCGCGAGTGCAGTCCATGCACAGC 180  
Qy 2047 CAGATTCTGGCGGCACCATGATCTGTCTATCGGGGGCATCATCTGGCCACGCTGCTG 2106  
Db 181 CAGATTCTGGCGGCACCATGATCTGTCTATCGGGGGCATCATCTGGCCACGCTGCTG 240  
Qy 2107 GTCTTCATCGTCATCTCATGTGCGCTACAGGCTCTCAACACAGAGGCCCCAGCAAG 2166  
Db 241 GTCTTCATCGTCATCTCATGTGCGCTACAGGCTCTCAACACAGAGGCCCCAGCAAG 300  
Qy 2167 ATGGCAGCGGCGTGGAGCAATGTACTCGCAGACCAACGCGGCCAGCACCGCTCCA 2226  
Db 301 ATGGCAGCGGCGTGGAGCAATGTACTCGCAGACCAACGCGGCCAGCACCGCTCCA 360  
Qy 2227 AGCAGCGCACAGCGGGGCGCCCGCGAGGGCCCGCGAAGGTGGTGGCAACGAG 2286  
Db 361 AGCAGCGCACAGCGGGGCGCCCGCGAGGGCCCGCGAAGGTGGTGGCAACGAG 420  
Qy 2287 CTCCTGGAGCTTACCGCGAGCTGCGCGCGCGAGTCTCTCTCTCCAGCTCCCTG 2346  
Db 421 CTCCTGGAGCTTACCGCGAGCTGCGCGCGCGAGTCTCTCTCTCCAGCTCCCTG 480  
Qy 2347 GGCAGTGGGAGGCTGCGGGGCTGGGAGCGGGCCCTTGGAGGATCCACCTCCGCCCCG 2406  
Db 481 GGCAGTGGGAGGCTGCGGGGCTGGGAGCGGGCCCTTGGAGGATCCACCTCCGCCCCG 540  
Qy 2407 CGCCCCAAGCCAGCCTTGTACCGCCTGATGGGGGCTTCGCTCTCTGTGACCTCAAGAT 2466  
Db 541 CGCCCCAAGCCAGCCTTGTACCGCCTGATGGGGGCTTCGCTCTCTGTGACCTCAAGAT 600  
Qy 2467 CAGAGAAAGGAGGAGCTCTGAGCTCCAGGACTCCAGCGGAGAGGGGCTGGGAGCTG 2526  
Db 601 CAGAGAAAGGAGGAGCTCTGAGCTCCAGGACTCCAGCGGAGAGGGGCTGGGAGCTG 660  
Qy 2527 GCCCGGGCGCACCTCGGACCGAGAGCCACTGTCTGGGGCCCTTGGCGCGCGGCGAGG 2586  
Db 661 GCCCGGGCGCACCTCGGACCGAGAGCCACTGTCTGGGGCCCTTGGCGCGCGGCGAGG 720  
Qy 2587 AGCCTGCTCCCTTCCGTTGGAGGGCAAGGCCAAACGAGCCACTCTCTTCGACATGGGG 2646

721 ACCCTGCTCCCTTCCGTTGAGGGCAAGCCAAACGACGACCTCTTCACATGGG 780  
2647 GACTTTGCTGCGCGGCGGAGGGTCTGCGCGGGGCTACAGTCTCTCGAAG 2706  
781 GACTTTGCTGCGCGGCGGAGGGTCTGCGCGGGGCTACAGTCTCTCGAAG 840  
2707 GTCTCGAATCTGACGAGCGGAGGCTCTCTGTCACGCGCATCTCTGCGCTTTGAG 2766  
841 GTCTCGAATCTGACGAGCGGAGGCTCTCTGTCACGCGCATCTCTGCGCTTTGAG 900  
2767 GAGAGTACCTGCTGGGGGCGCGGGGACTTTTGGCAGCTCCGAATGGTGTATGAGAGC 2826  
901 GAGAGTACCTGCTGGGGGCGCGGGGACTTTTGGCAGCTCCGAATGGTGTATGAGAGC 960  
2827 ACGGTCTAG 2835  
961 ACGGTCTAG 969

RESULT 6  
AAI60001  
ID AAI60001 standard; cDNA; 771 BP.  
XX  
AC AAI60001;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 3990.  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
OS Homo sapiens.  
XX  
WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
WPI: 2001-442253/47.  
DR P-PSDB; AAM40845.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
PS Claim 1; SEQ ID NO 3990; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities, chemotactic/chemokinetic activity, haemostatic  
CC Activin/inhibin activity, cancer diagnosis and therapy, drug screening,  
CC and thrombolytic activity, arthritis and inflammation, leukaemias and  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

XX Sequence 771 BP; 144 A; 259 C; 197 G; 171 T; 0 other;

Query Match 21.6%; Score 678.8; DB 22; Length 771;

Best Local Similarity 95.8%; Pred. No. 1.8e-133;

Matches 730; Conservative 0; Mismatches 27; Indels 5; Gaps 3;

QY 529 CCCAAGTACTGTCTGTCGCCAGAAATCTGTCTGAGTCACTGGGGACCTGTGCCCTCCAAG 588  
Db 8 CCCAAGACTGTCTGTCGCCAGAAATCTGTCTGAGTCACTGGGGACCTGTGCCCTCCAAG 67  
QY 589 GGGCTGCTCTTTTACCCCTGATATTGACCGCGGACAGTGGAGCTGCGCCTGGCGGC 648  
Db 68 GGGCTGCTCTTTTACCCCTGATATTGACCGCGGACAGTGGAGCTGCGCCTGGCGGC 127  
QY 649 AACTTCATCATCCACATCAGCCGCGGAGCTTTGCCAAATGAGCGGGCTGTGGAGCTG 708  
Db 128 AACTTCATCATCCACATCAGCCGCGGAGCTTTGCCAAATGAGCGGGCTGTGGAGCTG 187  
QY 709 ACCCTGTCAGGAACACCATCAGCCACATCAGCCCTTTTCTTCTGGACCTCGAGAGC 768  
Db 188 ACCCTGTCAGGAACACCATCAGCCACATCAGCCCTTTTCTTCTGGACCTCGAGAGC 247  
QY 769 CTCGCTCCCTGTCATCTTGACAGCAATCGGCTGCCAAGCCTTTGGGAGGACACCTCCGG 828  
Db 248 CTCGCTCCCTGTCATCTTGACAGCAATCGGCTGCCAAGCCTTTGGGAGGACACCTCCGG 307  
QY 829 GGCCTGGTCAACCTGCGAGCATTATCGTGAACAAACACAGCTGGCGGCGCATCGAGAT 888  
Db 308 GGCCTGGTCAACCTGCGAGCATTATCGTGAACAAACACAGCTGGCGGCGCATCGAGAT 367  
QY 889 GAGGCTTTTGGAGCTTCTGCTGACATTTGAGGATCTGGACCTCTCTCAACAACCTC 948  
Db 368 GAGGCTTTTGGAGCTTCTGCTGACATTTGAGGATCTGGACCTCTCTCAACAACCTC 427  
QY 949 CAT-GGCCTGCGCTGGAGTCCGT-GCGACGGATGG---TCAACCTCCACAGCTGAGCC 1003  
Db 428 CATGGGCTGCGCTGGAGTCCGTGGGACGCGATGGGTTCAACCTTCCACAGCTGAGCC 487  
QY 1004 TGGACCAACACCTGCTGGATCAGATCGCGGAGGCGACCTTTGCAGACCTCGAGAACTGG 1063  
Db 488 TGGACCAACACCTGCTGGATCAGATCGCGTGGAGGACCTTTGCAGACCTCGAGAACTGG 547  
QY 1064 CCGGCTGGATCTCACTCCAAATCGGCTGCGAAGCTGCCCTGATCCATCTTTGCC 1123  
Db 548 CCGGCTGGATCTCACTCCAAATCGGTTGCATAAGCTGCCCTGATCCATCTTTGCC 607  
QY 1124 GCTCCAGGCTTCGGCTTTGACGCCACACCTTTGCCCCACCTTGTCTCTTAGTTTGG 1183  
Db 608 GCTCCAGGCTTCGGCTTTGACGACACACCTTTGCCCCACCTTGTCTCTTAGTTTGG 667  
QY 1184 GGGGTAAACCCACTTCACTGCAATTTGTGAGCTTCTCTGGCTGGGAGGCTCGAGCGGAGC 1243  
Db 668 GGGGTAAACCCACTTCACTGCAATTTGTGAGCTTCTCTGGCTGGGAGGCTCGATCGGAGC 727  
QY 1244 ATGACCTGGAAACCTGTGGCTCCCGAGGGGCGCTCAAGGTC 1285  
Db 728 AAACCTTGGAAACTTGTGGTTCTCCAGTGGGCGCTCAAGGGTC 769

RESULT 7  
AAH25568

ID AAH25568 standard; DNA; 1887 BP.  
AC AAH25568;  
XX  
DT 05-SEP-2001 (first entry)  
XX  
DE Nucleotide sequence of an interferon omega-1 like protein NOV2.  
XX  
KW Interferon omega-1 like protein; Nov2; membrane bound protein;  
KW secreted protein; spermatogenesis; male infertility; neoplasia;  
KW blood circulation; immunological disorder; autoimmune disease;  
KW inflammatory disease; cardiovascular disease; metabolic disease;  
KW cancer; viral infection; acute lymphoblastic leukemia; glioma;  
KW neurological disease; neurodegenerative disorder; Alzheimer's disease;  
KW Parkinson's disease; hematopoietic disorder; neurodegenerative disorder;  
KW immune disorder; hematopoietic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 1..1887  
FT CDS /tag= a  
FT /product= "interferon omega-1 like protein NOV2"  
XX  
PN WO200142471-A2.  
XX  
PD 14-JUN-2001.  
XX  
XX 08-DEC-2000; 2000WO-US33463.  
XX  
XX 09-DEC-1999; 99US-0169887.  
PR 10-DEC-1999; 99US-0170230.  
PR 07-DEC-2000; 2000US-0170230.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Prayaga SK, Shinkets RA;  
PI WPI; 2001-381691/40.  
DR P-PSDB; AAB84469.  
XX  
XX Novel polypeptides designated as NOV polypeptides, useful in detection,  
PT prevention and treatment of a broad range of pathological states  
XX  
XX Claim 8; Page 15-16; 92pp; English.  
XX  
XX The present sequence encodes an interferon omega-1 like protein,  
CC designated NOV2. The protein is a membrane bound and secreted protein.  
CC NOV polypeptides and polynucleotides are useful for treating or  
CC preventing a NOV-associated disorder, and in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease such  
CC as NOV-associated disorder. They are also useful for treating disorders  
CC or syndromes including those involved in development, differentiation  
CC and activation of thymic immune cells, pathologies related to  
CC spermatogenesis and male infertility, diagnosis of several human  
CC neoplasias, in disease or pathologies of cells in blood circulation  
CC such as red blood cells and platelets, various immunological disorders  
CC and/or pathologies, autoimmune and inflammatory diseases,  
CC cardiovascular diseases, metabolic diseases, cancer growth and  
CC metastasis, viral infections, cancer therapy, acute lymphoblastic  
CC leukemia, in gliomas, neurological diseases, neurodegenerative disorders,  
CC Alzheimer's disease, Parkinson's disorder, and hematopoietic disorders.  
CC NOV polypeptides and polynucleotides are also useful for treating  
CC neurodegenerative disorders, immune disorders and hematopoietic  
CC disorders.  
XX  
SQ Sequence 1887 BP; 259 A; 736 C; 577 G; 315 T; 0 other;  
Query Match 20.2%; Score 634.6; DB 22; Length 1887;  
Best Local Similarity 61.3%; Pred. No. 4.6e-124;  
Matches 1079; Conservative 0; Mismatches 664; Indels 16; Gaps 3;  
507 GTTTCGGTGGTGGAGCGCTGCCCAAGTACTGTGTCTGCCAGAAATCTCTGTAGTCACT 566

Db 63 GTACGCCACACCCAGCCCATGTCCCGCGCGCTGCCAGACACAGTCTCGTCCCT 122  
Qy 567 GGGGACCCCTGTGCTCCAAAGGGGCTGCTCTTTGTACCCCTGATATTACCGGGGAC 626  
Db 123 AAGCGTGTGTGCTCCAGGGGAGGCTCTGTGTCGCCACCTCGCTGAGACCGCGGGC 182  
Qy 627 AGTGGAGTGGCGCTTGGGGGCAACTTTCATATCCACATCAGCCGCGAGACTTTGCCAA 686  
Db 183 AGCGAGTGTGGCTGGGAGACAACTTCATCGCTCGGTGCGCGCGGAGCTGGCCAA 242  
Qy 687 CATGAGGGGCTGGTGGACCTGTCCAGGAACACCATCAGCCCATCCAGCCCTT 746  
Db 243 CATGACAGCCCTGCTGCATCTGAGCTGTGCGGGAACACCATCCGCGACGTGGTCCGG 302  
Qy 747 TTCTCTTCTGGACCTCGAGAGCTCCGCTCCCTGCTATCTTGACAGCAATCGCTGCAAG 806  
Db 303 CGCCTTGGCGGACCTGGCGGCTCTGGTGCCCTGCACCTGGATGGCAACGGCTGACCTC 362  
Qy 807 CTTGGGGAGGACACCTTCGGGGCTTGGTCAACCTGCAGACCTTATCGTGAACAA 866  
Db 363 ACTGGGCGAGGCGCAGCTGCGCGCTTGGTCAACTTGGCCACCTCATCTCAGCAACA 422  
Qy 867 CCAGCTGGGCGCATCGCAGATGAGGCTTTTGGAGACTTCTCTGTGACATTTGAGGATCT 926  
Db 423 CCAGCTGGCAGCGCTGGCGGCGGCGCTGTGATGATTGTGCGGAGACACTGGAGACCT 482  
Qy 927 GGACCTCTCTACAAACCTCCATGGCTGCGGTGGGACTCCGTGCGAGCGATGTCGTC 986  
Db 483 CGACCTCTCTACAAACCTCGAGAGCTGCTGGAGGCTCTGGGGCGCTGGGCGG 542  
Qy 987 CCTCCACAGCTGAGCTGGACCAACCTGCTGGATCACATCGCGAGGCGACCTTTGC 1046  
Db 543 CGTCAACAGTTGGGCTCGACCAACCTGCTGGCTTCTGTGCGCGCGGCTTTTC 602  
Qy 1047 AGACTGAGAAACTGGCGCGCTGATCTCACTTCAATCGGCTGCGAGAGTGGCGCC 1106  
Db 603 CGCGCTGCACAAAGCTGGCGCGCTGACATGACCTTCAACCGCTGACCAATCCAC 662  
Qy 1107 TGATCCCATCTTTGCGCGCTCCAGCTTTCGGCTTTTGACAGCCACACCTTTTGGCC 1166  
Db 663 CGACCACTCTTCTCCGCTGCTCCCTGCTCGCAGGCCCCGGGGCTCGCCGCTCTGC 722  
Qy 1167 CTGTCTCTTTAGTTTGGGGTAAACCACTTCACTGCAATTTGTAGCTTCTCTGGCTGG 1226  
Db 723 CTGTGTGTGGCTTTTGGCGGAACCCCTGCACTGCAACTGCGAGTGTGTGGCTGCG 782  
Qy 1227 GAGGCTCGAGCGGAGATGACTTGAAACCTGTGGTCTCCCGAGGGGCTCAAGGGTGC 1286  
Db 783 TCGCTTGGCGGGAGGAGCACTCGAGGCTTGGCGTCCCCACCTGCTCTGGGGCGCG 842  
Qy 1287 CTACTTCTGGCATGTGCGTGAGGAGAGTTGTGTGGAGCGGCTCTCATCACCAGCA 1346  
Db 843 CTACTTCTGGCGGGTGGCGAGGAGTTTGTCTGCGAGGCGCGCGTGGTGGTCACTC 902  
Qy 1347 CACACACAAGTTGTGTTTCTGGAGGGCGAGCGGCACTCAAGTGAAGAGCAATTCG 1406  
Db 903 CTCACCACTCTGGCTGTGCGCGAGTTCGGCGGCTGCGCTGCGCTGCGGCGAGTGG 962  
Qy 1407 GGACCCCGAGCCCTTATCCACTGGGTAGCCCCGATGACCCCTGGTAGGAACTCTTC 1466  
Db 963 GGACCCAGAGCCCGTGTGCGTTGGTGTACCCCGAGGCGCGCTGTAGGCAACTCAAG 1022  
Qy 1467 AAGGACCGCTGTCTATGACATGGCACCTTGACATCTTCATCACCACATCTCAGACAG 1526  
Db 1023 CGGTGCCCGCGCTTCCCCAATGGGAGCGTGGAGTGTGTGTCACCGAGCGGGTGTGG 1082  
Qy 1527 TGGTGCCTTACCTGCTTGGCAATGCTGCCGAGAGGCGCAGGCCATGTTGGAGT 1586  
Db 1083 TGGCATCTTACCTGCTGCTGCGGCAATGACAGTGGGAGGCGCAGCTGCTGTGGAGT 1142  
Qy 1587 CTTCATCTGTCTAGCTGGCAGCTCAGCAACAGACACAGCCAGCCGCA--CTGACACCC 1644

Db	1143	GA	CTGTGGTCC	CCCCACACCTCTCTCAGCTAGCCACACACACCTGTGAC	CCCCCGCG	1202
Qy	1645	TC	CGCTCTCAG	CATCACTGGCTCCAGACAGACAGCCGGGAGGTGGAGGACAGTGGG		1704
Db	1203	GG	ACGGGATCT	GTGATGCTCTCACCCACCTCCCGCTGCTCTGTTCTGCCAAGTGGC		1262
Qy	1705	GG	CGGAGAGCCT	CCCCAAAGCCCCCGGAACGGGCTGTCTGTGTGTAAGTGACCAACC		1764
Db	1263	CG	ACACTGGG	CCCCCTA-----CCGACCTGGCGTCCAGTGACTGAGCACGGGGCC		1314
Qy	1765	AC	TCGGCGCTGGT	CAAGTGGTCTGTACACAGTACAGACCCCGGGTGAAGATGACACAG		1824
Db	1315	AC	AGCTGCTCT	GTTCAGTGGCGGATCAGCGGCTATCCCGGGCATCCGATGACACAG		1374
Qy	1825	CT	CCAGTACAACT	GTCTCAGCATGAGTACTGATTTACAGGATGATCCAGCCTCCAAAC		1884
Db	1375	AT	CCAGTACAA	CGCTCGGTGATGATCCTCTACAGGATGATCCCGGGGAGAGC		1434
Qy	1885	AA	GGCCTTCTGT	GTCAACAACTGGTGTACGGACTGGCTACGACTGTGTGTGTGGCC		1944
Db	1435	CG	CTCGTTCCT	GTGACGGACCTGGGCTCAGGCGGACCTACGATCTGTGCTGTGCGC		1494
Qy	1945	AT	TGGGATACAC	AGCAGCACTCAGGCCACCAACATCGTGGCTGGCGCCAGTTTC		2004
Db	1495	GT	GATGAGACAG	CGCCACGGGCTCAGGCCACCGGGCTGTGGCTGGCGCGCTTC		1554
Qy	2005	TT	CACCAAGGCT	GACTACCTACCGCAGTGCCAGTCCATGACAGCCAGATTTCTGGGCGCAC		2064
Db	1555	TC	CACCAAGCT	TGGCTGCGCCATGGGGCGCGCAGCTCCCTCTCTGGGCGGACG		1614
Qy	2065	AT	GATCTGTGTC	ATCGGGGATCATGTGGCCACGCTGTGCTCTTCATCGTCATCTTC		2124
Db	1615	AT	GATCATCGCT	TGGCGCGCTCATCTAGCTCGGTACTGTGCTTCATCTTCGTGCTG		1674
Qy	2125	AT	GTGGCTACAA	GGTCTGCAACACGAGCGCCCGCCAG-----CAGATGGGAGCGGCC		2178
Db	1675	CT	ATGCTGTACA	AGGTGCTGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGCT		1734
Qy	2179	GT	GAGCAATGTG	TACTCGAGACAAAGCGGCGCCAGCCGCTTCCAGCAGGCGACCA		2238
Db	1735	GT	TAGCAGCGTT	GTCTCCAGACCAAGCGGCGCCCTGGGCGCCACGCGCCCGCGCCCG		1794
Qy	2239	CG	CGGGGCGCC	CGCGAGG		2257
Db	1795	CC	CGCGCGCGG	AGCGCGG		1813
RESULT 8						
AAH78215						
ID	AAH78215 standard; DNA; 1887 BP.					
XX	AAH78215;					
AC	26-NOV-2001 (first entry)					
DE	Nucleotide sequence of a human secreted polypeptide.					
XX	Human; secreted polypeptide; nervous disease; muscular disease; tumour;					
KW	gastrointestinal ulceration; spinal cord disease; trachea disease;					
KW	thyroid gland disease; ovary disease; prostate disease; heart disease;					
KW	renal gland disease; small intestine disease; thymus disease;					
KW	lymph node disease; muscular system disease; colon disease;					
KW	lipase deficiency; cystic fibrosis; pancreatitis; clot formation;					
KW	myocardial infarction; angioplasty; liver disease; coagulation disorder;					
KW	microbial disease; immune disorder; inflammation; transplant rejection;					
KW	bone thickness; bone density; ferroxidase loss; apoptosis;					
KW	vascular smooth cell proliferation; vaccine; ss.					
XX	Homo sapiens.					
OS	Location/Qualifiers					
XX	1..1887					
FH	/*tag= a					
FT						
FT						

/product= "secreted polypeptide"	
XX	WO200166690-A2.
XX	13-SEP-2001.
XX	05-MAR-2001; 2001WO-US071143.
XX	06-MAR-2000; 2000US-0187107.
PR	13-MAR-2000; 2000US-0188916.
PR	03-OCT-2000; 2000US-0236874.
PR	03-OCT-2000; 2000US-0237846.
XX	(SMIK ) SMITHKLINE BEECHAM CORP.
PA	(SMIK ) SMITHKLINE BEECHAM PLC.
XX	Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
PI	WPI; 2001-570768/64.
XX	P-PSDB; AAG67523.
DR	Novel isolated secreted polypeptide useful for treating nervous and
XX	muscular diseases, gastrointestinal ulceration, coagulation and immune
PT	disorders, microbial diseases, inflammation and transplant rejection -
PT	Claim 2; Page 49-50; 102pp; English.
XX	The present sequence encodes a human secreted polypeptide. The
CC	secreted polypeptides and polynucleotides are useful for treating
CC	nervous and muscular diseases, for inhibiting tumour formation and
CC	metastasis, for treating gastrointestinal ulceration, for preventing
CC	and treating diseases in spinal cord, thyroid gland, ovary, prostate,
CC	renal gland, small intestine, heart, trachea, thymus, lymph node,
CC	muscular system and colon, for treating lipase deficiency in cystic
CC	fibrosis and pancreatitis, for treating undesirable clot formation
CC	such as myocardial infarction, during angioplasty and all surgical
CC	procedures that require decreased blood clot formation, for treating
CC	liver diseases, coagulation disorders and microbial diseases, for
CC	treating immune disorders, for treating inflammation and transplant
CC	rejection, for enhancing bone thickness and increasing bone density,
CC	for reducing the loss of essential ferroxidases, for suppressing
CC	apoptosis, and for regulating vascular smooth cell proliferation. They
CC	may also be used as vaccines.
XX	Sequence 1887 BP; 259 A; 736 C; 577 G; 315 T; 0 other;
XX	Query Match 20.2%; Score 634.6; DB 22; Length 1887;
XX	Best Local Similarity 61.3%; Pred. No. 4.6e-124;
XX	Matches 1079; Conservative 0; Mismatches 664; Indels 16; Gaps 3;
Qy	507 GTTTGCGGTGGTGCAGCGCTGCCCAAGTACTGTCTGCCAGAACTCTGTGAGTCACT 566
Db	63 GTCAGGCCACACCCAGGCGCCATGTCCCGCGCTGCCGTGCCAGACACAGTCGCTGCCCT 122
Qy	567 GGGGACCCCTGTGCCCTCCAAAGGGGTGCTCTTTGTACCCCTGATATTGACCGCGGAC 626
Db	123 AAGCGTCTGTGCCCGAGGGGAGGCTCTCTGTTCGTGCCACCTCGCTGGACCGCGGGC 182
Qy	627 AGTGGAGCTGGCGCTGGCGGCGCAACTTCATCATCCACATCAGCGCCAGGACTTTGCCAA 686
Db	183 AGCCGAGCTGGCGTGGCGAGCAAACTTCATCGCTCGTGGCGCGCGGACCTGGCCAA 242
Qy	687 CATGACGGGCTGTGTGACCTGACCTGTCCAGGAACCATCAGCCACATCCAGCCCTT 746
Db	243 CATGACAGGCGCTGTGTCATCTGAGCCCTGTCCGGAAACCATCCGACAGTGGTGCGGG 302
Qy	747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
Db	303 CGCCTTTCGCCGACCTGCGGGCCCTGCGTGCCTGCACCTGGATGGCAACCGGTGACCTC 362
Qy	807 CTTTGGGAGGACACCTTCGGGGCTGGTCAACTGCGAGCACCTTATCGTGAAACAA 866
Db	363 ACTGGCGAGGGCCAGCTGCGCGGCTGGTCAACTTGGCGCACCTCATCTCAGCAACAA 422



QY 867 CCAGCTGGCGGCATCGCAGATGAGGCTTTTGGAGACTTCTCTGCTGACATGAGGATCT 926  
Db |||||  
QY 927 GGACCTCTCTACACAACTTCATGCGCTGCGGTGGGACTCCGTGCGACGATGGTCAA 986  
Db |||||  
QY 987 CCTCACAGCTGAGCTGGACCAACAACTGTGTGATCACATCGCGGAGGGACCTTTGTC 1046  
Db |||||  
QY 1047 AGACTGCGAAACTGGCGCGCTGATCTCACTTCAATCGGCTGCGAAGCTCCCGCC 1106  
Db |||||  
QY 1107 TGATCCCATTTGCGCGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTTGGCCACCC 1166  
Db |||||  
QY 1167 CTGTGCTTTAGTTTGGGGTAAACCACTTCACTGCAATTGTGAGCTTCTCTGGCTGG 1226  
Db |||||  
QY 1227 GAGGCTCGAGCGGACGATGACTCGAAACCTGTGCTCCCGAGGGGCTTCAAGGGTGG 1286  
Db |||||  
QY 1287 CTACTTCTGGCATGTGCTGAGAGAGAGTTTGTGCGAGCGCGCTCTCATCACCAGCA 1346  
Db |||||  
QY 1347 CACACAAAGTTGCTGTCTGAGGGCCAGCGGCCACACCAAGTCAAGTCAAGGCCATTGG 1406  
Db |||||  
QY 1407 GGACCCCGAGCCCTTATCCACTGGGTAGCCCCCGATGACCCGCTGGTAGGAACTCTCTC 1466  
Db |||||  
QY 1467 AAGGACCGTGTCTATGACAAATGGACCCCTGACATCTTCATCACCACATCTCAGACAG 1526  
Db |||||  
QY 1527 TGGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1586  
Db |||||  
QY 1587 CTTCATGCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644  
Db |||||  
QY 1645 TCCGCGCTCTAGACATCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1704  
Db |||||  
QY 1705 GCGGAGAGCTTCCCAAAAGCCCCCGGAAACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764  
Db |||||  
QY 1765 ACTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824  
Db |||||  
QY 1825 CTGCGAGTACATGCTCTGACATGAGTACTGATTTACAGATGATCCAGCCTTCAAC 1884  
Db |||||  
QY 1885 AAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944  
Db |||||  
QY 1945 ATGTGGGATGACACAGCCACGACACTACGGCCACCAACATCGTGGGCTGCGCCGAGTTTC 2004  
Db |||||  
QY 2005 TTCCACCAAGGCTGACTACCCGAGTGCCAGTCCATGACACAGCAGATTCTGGGCGGCACC 2064  
Db |||||  
QY 2065 ATGATCTCTGGTTCATCGGGGGCATCATCGTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTG 2124  
Db |||||  
QY 2125 ATGTGCGCTACAGGCTGTGCAACACAGAGGCCCCAG-----CAAGATGCGAGGGGCC 2178  
Db |||||  
QY 2179 GTGACCAATGTACTCGCAGACCAACGCGGCCACCGCCTCCAGAGCGCACCA 2238  
Db |||||  
QY 2239 GCGGGGCGCCCGCCGAGG 2257  
Db |||||  
QY 2257 GCGGGGCGCCCGCCGAGG 2257  
Db |||||

RESULT 9

AAI66986

ID AAI66986 standard; cDNA; 1887 BP.

AC AAI66986;

DT 30-JAN-2002 (first entry)

XX Human LRR family member, 33395 coding sequence.

XX LRR; leucine rich repeat; 33395; cytostatic; anti-HIV; antidiabetic;  
XX antiarthritic; neuroprotective; dermatological; immunosuppressive; ss;  
XX antiinflammatory; antiasthmatic; antiulcer; antianaemic; hepatotropic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1887

FT FT /\*tag= a

FT FT /product= "33395"

XX WO200172827-A2.

XX 04-OCT-2001.

XX 23-MAR-2001; 2001WO-US09470.

XX 24-MAR-2000; 2000US-191863P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Glucksmann MA;

XX WPI; 2001-626254/72.

XX P-PSDB; AAG65805.

XX New polypeptide for preventing or treating disorders associated with  
XX cellular adhesion, proliferation or differentiation, comprises  
XX polypeptide 33395, a member of the leucine rich repeat protein family

XX Claim 1; Fig 1; 133pp; English.

XX The invention provides an isolated nucleic acid encoding a polypeptide  
XX of the leucine rich repeat (LRR) family, designated 33395. The 33395  
XX polypeptide can be expressed by standard recombinant methodology. The  
XX 33395 polynucleotides and polypeptide can be used to prevent or treat  
XX disorders associated with 33395 expression, for example those involving



CC aberrant cellular adhesion, proliferation or differentiation. Specific.  
 CC examples include melanoma, juvenile AIDS, diabetes mellitus,  
 CC rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis,  
 CC Chron's disease, ulcerative colitis, asthma, anemia, and chronic active  
 CC hepatitis. The present sequence represents the coding sequence of the  
 CC human 33395 polypeptide.

XX  
 SQ Sequence 1887 BP; 257 A; 735 C; 579 G; 316 T; 0 other;

Query Match 20.1%; Score 633; DB 22; Length 1887;

Best Local Similarity 61.3%; Pred. No. 1e-123;

Matches 1078; Conservative 0; Mismatches 665; Indels 16; Gaps 3;

QY 507 GTTTGCGGTGTCGACGCTGCGCCCAAGTACTGTGTCTGCCAGATCTGTCTGATCACT 566  
 DB 63 GTACGACACACGAGCCCATGTCCTCCCGCGCTGCCGTGCCAGACACAGTCTGCCCT 122  
 QY 567 GGGGACCTGTGCGCCCTCAAGGGGTGCTCTTTGACCCCTGATATTACCGGGAC 626  
 DB 123 AAGCGTGTGTGCCAGGGGAGGCTCTCTGTGTGTCGACCATCTGCTGGACCGGGC 182  
 QY 627 AGTGGAGCTGCGCTGCGGCGCAACTTCATCTCATCCACATCAGCGCCAGGACTTTGCCAA 686  
 DB 183 AGCGAGCTGCGCTGGCAGACAACTTCATGCGCTCGGTGCGCGCGGCACTGGCCAA 242  
 QY 687 CATGACGGGCTGTGACCTGACCTCTGTCAGGAACACCATCAGCCACATCCAGCCCTT 746  
 DB 243 CATGACAGGCTGTGATCTGAGCTGTGCGGAAACACCATCCGCGACGTGCTGCCGG 302  
 QY 747 TTCTTCTGAGCTGAGAGCTGCGCTCTCTGATCTTGACAGCAATCGGTGCCAAG 806  
 DB 303 CGCTTTCGCGACCTGCGGCGCTGTGCTGCTGACCTGATGGGAAACCGGTGACCTC 362  
 QY 807 CTTTGGGAGGACACCTTCGCGGCTGTGTCACCTGACGACCTTATCGTGAACAA 866  
 DB 363 ACTGGGCGAGGCGAGCTGCGGCGCTGTGTCACCTGCGGACCTATCTCAGCAACAA 422  
 QY 867 CCAGCTGGCGGATCCAGATGAGGCTTTTGAGGACTTCCTGTGACATTTGAGGATCT 926  
 DB 423 CCAGCTGGAGCGCTGCGGCGCGGCTGTGATGATGTTGCCGAGACACTGGAGGACCT 482  
 QY 927 GGACTCTCTTACAAACCTTCCATGCGCTGCGGACTTCCGTCGAGCGATGTGTCAA 986  
 DB 483 CGACTCTCTTACAAACCTTCCATGCGGACTTCCGTCGAGCGCTGCGGCGCTTTC 542  
 QY 987 CCTCACAGCTGAGCTGAGCCACCACTGCTGATCACATCCGCGAGGCGACTTTGC 1046  
 DB 543 CGTCAACAGCTTGGCGCTTCGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602  
 QY 1047 AGACTCGAGAACTGGCGCGCTTGGATCTCACTTCCAAATCGGCTGAGAGCTGCCCC 1106  
 DB 603 CGGCTGTCAGAGCTGGCGCGCTGAGATGATGACCTTCCAAACGCTGACCAAAATCCCA 662  
 QY 1107 TGATCCCATCTTTGCGCGCTTCCAGGCTTCCGCTTTGACGACACACCTTTGCCCA 1166  
 DB 663 CGACCCACTCTTCTCCGCGCTGCGCTGCTGCGGAGGCGCGGCTGCGCGCTTTC 722  
 QY 1167 CTTGTCTTTAGTTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTTGGCTGCG 1226  
 DB 723 CTTGTGCTTGGCTTTGGCGGAAACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 782  
 QY 1227 GAGGCTCGAGCGGAGGATGACCTGGAACCTGTGGCTTCCCGAGGGGCTCAAGGGTCG 1286  
 DB 783 TCGCTTGGCGCGGAGGAGGACCTCGAGGCTGCGGCTGCCCACTGCTGCTGCGGCGCG 842  
 QY 1287 CTACTTCTGGCATGCTGCTGAGGAGGATTTGCTGCGAGCGCGCTCTCATCACCAGCA 1346  
 DB 843 CTACTTTTGGCGGTGGCGAGGAGTTTGTCTGCGAGCGCGCTGCTGCTGCTGCTGCT 902  
 QY 1347 CACACAAAGTTGCTGCTTCTGAGGGCGGAGGCGGCACTCAAGTGCAAGGCAATGG 1406  
 DB 903 CTACCCACCTTGGCTGTGCGCGAGTGGCGGCTGCGCTGCGCTGCGGCGGAGTGGG 962

QY 1407 GGACCCAGCCCGCTTATCCATGCTGCCAATGTCGCGAGAGGCGCAACCGCATGTTGAGGT 1466  
 DB 963 GGACCCAGAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022  
 QY 1467 AAGGACCGTGTCTATGACAAATGGACCTGTCGACATCTTTCATCACCACATCTCAGACAG 1526  
 DB 1023 CCGTGCCTGCGCTTCCCAATGGGACGCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1082  
 QY 1527 TGTGCTTCCATGCTGCTGCCAATGTCGCGAGAGGCGCAACCGCATGTTGAGGT 1586  
 DB 1083 TGGCATCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142  
 QY 1587 CTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644  
 DB 1143 GACTGTGGTTCCTCCACCTCTCTCAGCTGAGCAACAGCAGCTGCTGCTGCTGCTGCTGCT 1202  
 QY 1645 TCCGCTCTCAGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704  
 DB 1203 GGACGGGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262  
 QY 1705 GGCGGAGAGCTCTCCAAAGCCCGGAAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764  
 DB 1263 CGACACTGGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314  
 QY 1765 ACCTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1824  
 DB 1315 ACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1374  
 QY 1825 CTGAGTACAACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1884  
 DB 1375 ATCAGTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1434  
 QY 1885 AAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944  
 DB 1435 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1494  
 QY 1945 ATGTGGGATGACAGACGACGACACTCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2004  
 DB 1495 GTGTATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1554  
 QY 2005 TTCCACAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2064  
 DB 1555 TCCACGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1614  
 QY 2065 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2124  
 DB 1615 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1674  
 QY 2125 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2178  
 DB 1675 CTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1734  
 QY 2179 GTGAGCAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2238  
 DB 1735 GTTAGCAGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1794  
 QY 2239 GCGGCGGCGCGCGCGCG 2257  
 DB 1795 CCGCGCGCGAGCGCGCG 1813

## RESULT 10

AAI66985

ID AAI66985 standard; cDNA; 2637 BP.

XX

AC AAI66985;

XX

DT 30-JAN-2002 (first entry)

XX Human leucine-rich repeat (LRR) family member, 33395 cDNA sequence.

DE

XX LRR; leucine rich repeat; 33395; cytosolic; anti-HIV; antidiabetic;

XX antiarthritic; neuroprotective; dermatological; immunosuppressive; ss;

KW



Novel proteins and polynucleotides of secreted proteins useful for treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis, diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury

Claim 1; Page 124-127; 145pp; English.

The invention relates to human secreted polypeptides designated LP095, LP194, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic acid molecules encoding such polypeptides. Novel secreted proteins of the invention are used for treating diseases such as arteriosclerosis, Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis, arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia, reperfusion injury, neoplasms and cancer especially liver cancer. They are also used for wound healing. Polynucleotides of the invention can be used to generate transgenic animals or knock out animals, which in turn, are useful in the development and screening of therapeutically useful reagents for use in the treatment of diseases associated with LP polypeptide associated activity. They are also used in gene therapy. The present sequence is human LP220 secreted protein encoding cDNA. LP220 gene is located on chromosome 11q13.

PT Novel proteins and polynucleotides of secreted proteins useful for  
PT treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,  
PT diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury  
PS  
XX  
XX  
CC The invention relates to human secreted polypeptides designated LP095,  
CC LP194, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic  
CC acid molecules encoding such polypeptides. Novel secreted proteins of  
CC the invention are used for treating diseases such as arteriosclerosis,  
CC Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,  
CC arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe  
CC combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia,  
CC reperfusion injury, neoplasms and cancer especially liver cancer. They  
CC are also used for wound healing. Polynucleotides of the invention can  
CC be used to generate transgenic animals or knock out animals, which in  
CC turn, are useful in the development and screening of therapeutically  
CC useful reagents for use in the treatment of diseases associated with  
CC LP polypeptide associated activity. They are also used in gene therapy.  
CC The present sequence is human LP220 secreted protein encoding cDNA.  
CC LP220 gene is located on chromosome 11q13.  
XX  
SQ Sequence 2369 BP; 342 A; 855 C; 755 G; 417 T; 0 other;

Query Match 19.8%; Score 622.6; DB 24; Length 2369;  
Best Local Similarity 62.3%; Pred. No. 1.6e-121;  
Matches 1089; Conservative 0; Mismatches 609; Indels 51; Gaps 5;

QY 520 GACGCTCCCAAGTACTGTCTGCCAGAACTGTCTGTGAGTCACTGGGACCTGTGTC 579  
DB 209 GCCGCTCCGCGCTGCCCTGGCTGTCCAGAACTGTCCGAGTCTCAGACCTCTGT 268  
QY 580 CCCTCCAGGGGCTGCTTTTGTACCCCTGTATTGACCGGGGACAGTGGAGCTCGC 639  
DB 269 GCCACCGAGGCTGCTTTGTGCGGCCCAACGTGACCGGCGCACAGTGGAGCTGG 328  
QY 640 CTGGGCGCAACTTCACTCATCCATCAGCCGCGCAGGACTTTGCAACATCAGCGGGCTG 699  
DB 329 CTGGGTGACAACTTCACTCAGGCTTGGGGCCCTGACTTCCGACATCAGCGGACTG 388  
QY 700 GTGGACCTGACCTGTCTCCAGGAACACCATCAGCCACATCCAGCCCTTTCTTTCTGGAC 759  
DB 389 GTGGACCTGACACTGTCTCCGAAATGCCATCACCAGTGGGGCGCCGCTTTGGGGAC 448  
QY 760 CTCGAGAGCTCCGCTCCCTGCACTTTGACAGCATCGCTGCCAAGCTTTGGGAGGAC 819  
DB 449 CTCGAGAGCTCCGCTCCCTGCACTTTGACAGCATCGCTGCCAAGCTTTGGGAGGAC 508  
QY 820 ACCCTCCGGGCTCGTCAACCTGCGACACCTTATCGTGAACAAACAGCTGGGCGGC 879  
DB 509 AGCCTCCGGGCGCCGCTCAATCTGCGACACCTCATCTCCAGCGGCAACAGCTGGGCGGC 568  
QY 880 ATCGCAGATGAGGCTTTGAGGACTTCTGCTGACATTTGGAGGATCTGGACCTCTCTAC 939  
DB 569 ATCGCGCGGGAGGCTTTCGACGACTTCTTAGAGAGCCTGGAGGACTGGACCTGTCTAC 628  
QY 940 AACAACTCCATGGCTCCGCTGGGACTCCGCTGCGACGATGCTCAACCTCCACAGCTG 999  
DB 629 AACAACTCCGAGGTGCCCTGGGCGGCTATCGGCGCATCGCTGCCCTGCACACCTTC 688  
QY 1000 AGCCTGGACCAACCTGCTGGATCATATCGCGAGGCGACCTTTGACAGCTTCAGAAA 1059  
DB 689 AACCTGGAACCACTTATTTAGCGCACTGCCCCAGGCGCTTCCGCCAGCTCGGTGAG 748  
QY 1060 CTGGCGCGCTGTGATCTCACTCCCAATCGGCTGCGAAGCTGCCCTCTGATCCCTTT 1119  
DB 749 CTCTCCGCTGGACCTCACTCCAAACCGCTGGCCAGCTGGCTCGGACCGCTTTTC 808  
QY 1120 GCCCGCTCCAGGCTTCGGCTTTGACAGCGCACACCTTTGCCCCACCTTGTCTTAGT 1179  
DB 809 TCTCTGGGCTGTAT-----GCAGAGGCTCTCCCGCCCCCTGGTGTGAGC 856

RESULT 11  
ID AAD38695 standard; cDNA; 2369 BP.  
XX AAD38695;  
AC AAD38695;  
XX 23-SEP-2002 (first entry)  
DE Human LP220 secreted protein encoding cDNA.  
XX Human; secreted protein; arteriosclerosis; Alzheimer's disease; LP220;  
XX diabetic retinopathy; severe combined immunodeficiency; pancreatitis;  
KW rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;  
KW reperfusion injury; arteriosclerosis; wound healing; transgenic animal;  
KW gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;  
KW chromosome 11q13; gene; ss.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH 167..2074  
FT CDS /\*tag= a  
FT /\*product= "Human LP220 secreted protein"  
FT sig\_peptide 167..214  
FT /\*tag= b  
FT mat\_peptide 215..2071  
FT /\*tag= c  
FT /\*product= "Mature human LP220 secreted protein"  
XX WO200226801-A2.  
XX 04-APR-2002.  
XX 14-SEP-2001; 2001WO-US26026.  
XX 28-SEP-2000; 2000US-236088P.  
XX (ELIL ) LILLY & CO ELI.  
XX Su EW, Wang H;  
XX WPI; 2002-471259/50.  
DR P-PSDB; .AAE23980.  
XX

QY 1180 TTTGGGGTAACCCACTGCAATTTGTGAGCTTCTTGCTGCGGAGGCTCAGCGG 1239  
DB |||||  
857 TTTAGCGGAACCCCTGCACTGCAATTTGTGAGCTTCTTGCTGCGGAGGCTCAGCGG 916  
QY 1240 GACGATGACCTGGAACCTGCTGCTCCCGAGGGGCTCAAGGCTCGCTACTTCTGGCAT 1299  
DB |||||  
917 CCGGAGCAGCTGGAACCTGCTGCTCCCGAGGGGCTCAAGGCTCGCTACTTCTGGCAT 976  
QY 1300 GTGCGTAGGAGAGTTTGTGCGAGCGGCTCTCATCACCAGCAGCAGCAGCAAGTTG 1359  
DB |||||  
977 GTGCGTAGGAGAGTTTGTGCGAGCGGCTCTCATCACCAGCAGCAGCAGCAGCCTC 1036  
QY 1360 CTGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1419  
DB |||||  
1037 TGGGCTGCTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1096  
QY 1420 CTATTCACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1479  
DB |||||  
1097 ACCATGCACTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1156  
QY 1480 TATGACAAATGGCAGCCTGGACATCTTATCACCACATCTCAGGACAGTGGTGGTGGTGG 1539  
DB |||||  
1157 TTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1216  
QY 1540 TGCAATGCTGCAATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1599  
DB |||||  
1217 TGCAATGCTGCAATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1276  
QY 1600 CTGCGACACCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1659  
DB |||||  
1277 TTGCCCCCATGTTGGGAACAGCAGTCCCGAGGGGGGCGGCC-----CGGGGCGCTCGGAC 1330  
QY 1660 ATCACTGGCTCCAGCAAGACACCGGGAGGTGGAGGAGTGGAGGAGTGGAGGAGTGGAG 1719  
DB |||||  
1331 ATCGCGGCTCGCTCGCACTGCTGCGAGGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAG 1386  
QY 1720 AAAAGCCCCCGGAAAGCGGCTGTGTTGTCTGAAGTGAACCAACCTCGGCGCTCGTGC 1779  
DB |||||  
1387 -----AGCCGTGAGGTGACGGAGGTGACCGGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAG 1429  
QY 1780 AAGTGGTCTGTGAGCAAGTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1839  
DB |||||  
1430 AGCTGGGTCCTCGGCGGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1489  
QY 1840 TCTGACGATGAGTACTGATTTACAGGATGATCCAGGCTCCAAAGAGCCTTCGTGGTC 1899  
DB |||||  
1490 ACGGAGATGACACCTCATCTACCGGATGTCAGGCTGTCAGGCTGTCAGGCTGTCAGGCTG 1549  
QY 1900 AACAACCTGGTGTGAGGAGTGGCTACGACTTGTGTGCTGCGGAGTGGGATGAGGATGACACA 1959  
DB |||||  
1550 AAGCACCTGTCCCGGCGCTGACTATGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1609  
QY 1960 G---CCACGACACTACGCGGACCAACATCTGGGCTGCGGCGGAGTCTTCCACAGGCT 2016  
DB |||||  
1610 GGGGCTCTGACTCTACGCGGACCAAGGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCTG 1669  
QY 2017 GACTACCCGAGTGGCAGTCCATGACAGCAGATTTCTGGGCGGACCATGATCTCTGTC 2076  
DB |||||  
1670 GCTGCGGCTGTGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1729  
QY 2077 ATCGGGGGCATCATGTCGCGGAGTGTGCTTCTCATGCTCATCTCATGTCGTCGCTAC 2136  
DB |||||  
1730 GTGGGGGTGTGCTGGTGGCTGCTTACTGGTCTTCTCATGTCGTCGCTTCTGCTGGGCTG 1789  
QY 2137 AAGTGTGCAACAGCAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2196  
DB |||||  
1790 CGGGGGGCGGAAATGGCGGCTCTCCGCTCAAG-----CTCAGCCACGCTCCAGTCC 1840  
QY 2197 CAGACCAAGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2256  
DB |||||  
1841 CAGACCAATGAGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1900  
QY 2257 GGCGCGCGCG 2265

Db 1901 CGGCGCAG 1909

RESULT 12  
ABK70010

ID ABK70010 standard; DNA; 2855 BP.

XX ABK70010;

AC 15-JUL-2002 (first entry)

DT cDNA encoding human Pro peptide #50.

DE Human; ss; gene; PRO; secreted protein; transmembrane protein;  
genetic disorder; tumour; cancer.

XX Homo sapiens.

PN WO200224888-A2.

PD 28-MAR-2002.

PF 29-AUG-2001; 2001WO-US27099.

PR 01-SEP-2000; 2000US-229896P.

PR 05-SEP-2000; 2000US-230621P.

PR 22-SEP-2000; 2000US-235147P.

PR 10-NOV-2000; 2000WO-US30873.

PR 12-JAN-2001; 2001US-261878P.

PR 16-JAN-2001; 2001US-261910P.

PR 16-JAN-2001; 2001US-262150P.

PR 25-JAN-2001; 2001US-264395P.

PR 02-FEB-2001; 2001US-266421P.

PR 09-FEB-2001; 2001US-267623P.

PR 28-FEB-2001; 2001WO-US06520.

PR 09-MAR-2001; 2001US-274399P.

PR 03-APR-2001; 2001US-280982P.

PR 04-APR-2001; 2001US-282129P.

PR 09-MAY-2001; 2001US-290589P.

PR 25-MAY-2001; 2001WO-US17092.

PR 01-JUN-2001; 2001WO-US17800.

PR 20-JUN-2001; 2001WO-US19692.

PR 29-JUN-2001; 2001WO-US21066.

PR 09-JUL-2001; 2001WO-US21735.

XX (GETH ) GENENTECH INC.

XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;

XX Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;

XX Fong S;

XX WPI; 2002-362426/39.

XX P-PSDB; ABG34079.

XX New PRO polypeptides and polynucleotides encoding the polypeptides,  
useful in gene therapy, chromosome identification, tissue typing, or  
for genetic analysis of individuals with genetic disorders

Claim 2; Figure 99; 218pp; English.

This invention relates to the cDNA and protein sequences of novel  
secreted and transmembrane polypeptides PRO polypeptides. The  
invention also comprises a method for producing the proteins of the  
invention by recombinant means and antibodies specific for the protein  
of the invention. The antibody may be used for detecting the PRO  
proteins of the invention and may be used to modify their activity.  
Polynucleotides may be used as hybridisation probes for a cDNA library  
to isolate the full-length PRO cDNA or to isolate other cDNAs, to  
construct hybridisation probes for mapping the gene which encodes that  
PRO and for genetic analysis of individuals with genetic disorders, in



DE Nucleotide sequence of a human secreted polypeptide.

XX Human; secreted polypeptide; nervous disease; muscular disease; tumour;  
 KW gastrointestinal ulceration; spinal cord disease; trachea disease;  
 KW thyroid gland disease; ovary disease; prostate disease; heart disease;  
 KW renal gland disease; small intestine disease; thymus disease;  
 KW lymph node disease; muscular system disease; colon disease;  
 KW lipase deficiency; cystic fibrosis; pancreatitis; clot formation;  
 KW myocardial infarction; angioplasty; liver disease; coagulation disorder;  
 KW microbial disease; immune disorder; inflammation; transplant rejection;  
 KW bone thickness; bone density; ferroxidase loss; apoptosis;  
 KW vascular smooth cell proliferation; vaccine; ss.

XX Homo sapiens.

OS

XX

FH Key Location/Qualifiers

FT 1..2316

FT CDS /tag= a

FT /product= "secreted polypeptide"

XX

XX WO200166690-A2.

XX 13-SEP-2001.

XX

XX 05-MAR-2001; 2001WO-US07143.

XX

XX 06-MAR-2000; 2000US-0187107.

PR 13-MAR-2000; 2000US-0188916.

PR 03-OCT-2000; 2000US-0236874.

PR 03-OCT-2000; 2000US-0237846.

XX

XX (SMIK ) SMITHKLINE BEECHAM CORP.

PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX

XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;

XX

XX WPI; 2001-570768/64.

DR P-PSDB; AAG67512.

XX

XX Novel isolated secreted polypeptide useful for treating nervous and  
 PT muscular diseases, gastrointestinal ulceration, coagulation and immune  
 PT disorders, microbial diseases, inflammation and transplant rejection -

XX

PS Claim 2; Page 43-44; 102pp; English.

XX

XX The present sequence encodes a human secreted polypeptide. The  
 CC secreted polypeptides and polynucleotides are useful for treating  
 CC nervous and muscular diseases, for inhibiting tumour formation and  
 CC metastasis, for treating gastrointestinal ulceration, for preventing  
 CC and treating diseases in spinal cord, thyroid gland, ovary, prostate,  
 CC renal gland, small intestine, heart, trachea, thymus, lymph node,  
 CC muscular system and colon, for treating lipase deficiency in cystic  
 CC fibrosis and pancreatitis, for treating undesirable clot formation  
 CC such as myocardial infarction, during angioplasty and all surgical  
 CC procedures that require decreased blood clot formation, for treating  
 CC liver diseases, coagulation disorders and microbial diseases, for  
 CC treating immune disorders, for treating inflammation and transplant  
 CC rejection, for enhancing bone thickness and increasing bone density,  
 CC for reducing the loss of essential ferroxidases, for suppressing  
 CC apoptosis, and for regulating vascular smooth cell proliferation. They  
 CC may also be used as vaccines.

XX

XX Sequence 2316 BP; 347 A; 836 C; 758 G; 375 T; 0 other;

XX

XX Query Match 19.7%; Score 618.8; DB 22; Length 2316;  
 Best Local Similarity 60.7%; Pred. No. 1e-120;  
 Matches 1129; Conservative 0; Mismatches 677; Indels 54; Gaps 5;

XX

XX 487 CTGCTAGCGTTTGGCATGGCTTTCGGTGGTGGCGCGCTGCCCAAGTACTGTGCTGC 546  
 DB 61 CTGCTGCTGCTTGGCGGGGGGATCTCGTGGCCAGCCCTCGCCCGCGCGCTGCATCTGC 120

XX

XX 547 CAGAATCTGTCTAGTCACTGGGGACCCCTGTGCCCTCCAAAGGGGTGCTCTTTGTACCC 606

121 CAGAACGTTGGCGCCACACACTGACAAATGCTGTGGCGCAAGACCGGTTCGCTCTTTGTCGG 180

607 CCGTATATTACCGCGGACAGTGTGAGCTGTGGCGGCAAACTTATCATCCACATC 666

181 CCGCGCATCGACCGCGGCGTGTGAGCTGTGGCTACCGCAAACTTATCGCGCGCGTG 240

667 AGCCGCCAGGACTTTGCCAAATGACGGGCTGTGGACCTGACCTGTGCCAGGAACACC 726

241 CGCCGCCGAGACTTCGCCAAATGACAGCTGTGGTGCACCTCACCTCTCTCCCGGAACCC 300

727 ATCAGCCACATCCAGCCCTTTTCTCTTGTGGACCTCGAGAGCTCGGCTCCCTGCATCTT 786

301 ATCGGCCAGGTGGCAGCTGGCGCTTCGCCGACCTCGCTGTGGCGCTCGGCGCTGCACCTG 360

787 GACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCTCCCGGGGCTGTGTCAACCTGCAG 846

361 GACAGCAACCGCTGGCGGAGGTGCGCGGACAGCTCCGCGGCTGTGGCAACCTCCGC 420

847 CACCTTATCGTGAAACAAACAGCTGGCGGCGATCGCAGATGAGGCTTTTGAGGACTTC 906

421 CACCTGATCTTTGAAACAAACAGATCGCGCGGTGGAGTTCGGCGGCTTTTACGCGCTTC 480

907 CTGCTGACATTTGGAGGATCTGCACCTCTCTTACAAACCTCCATGGCTGCGGTGGGAC 966

481 CTGTCCACCGTGGAGGACCTGGATCTGTCTTAAACAACTGGAGGCGCTTGCCTGGGAG 540

967 TCCGTGCGACGATGCTCAACCTCCACAGCTGTGAGCTGTGACCAACCTGTGTGATCAC 1026

541 GCGTGGCGCAGATGTAACCTTAAACCTTACGCTGGACCAACCTCATTCGACCAAC 600

1027 ATCGCGGAGGCGACCTTTTCAGACCTGCAGAACTGGCGCGCTGGATCTCACTCTCAAT 1086

601 ATCGCGGAGGCGACCTTCTGTGAGCTTCAAGCTGTGCGCTGTGACATGACCTTCCAAC 660

1087 CGGCTGCGAAGCTGCGCGCTGATCCCATCTTTTGGCGCTCCAGGCTTCGGCTTTTGACA 1146

661 CGCTTGCATAAACTCCCGCGCGAGGCTTCTTCTGAGGTGCGAGG-----GCACC 711

1147 GCCACACCTTTTGGCGCGCGCTTGTCTTGTAGTTTGGGGTAAACCTTCACTCTCAAT 1206

712 GGGCCAAAGCGCGCGCGCTGACCTTGTGCGCGCAACCTTGTGAGCTGACTGCAAC 771

1207 TGTGAGCTTCTGTGCTGGGAGGCTCGAGCGGAGCATGACCTTGAAACCTGTGCTCC 1266

772 TCGGAGCTGCTCTGTGCTGGCGGCTGACCCCGAGGAGCATTAGAGACCTTGCGCACG 831

1267 CCAGGGGCGCTCAAGGCTGCTTCTTGTGATGTGCGTGTGAGGAGAGTTTGTGCGAG 1326

832 CCGCAACACCTCACCGACCGCTTCTTGTGCTCCATCCCGAGGAGAGTTTCTGTGTGAG 891

1327 CGGCTCTCATCACCCAGCACAC---ACACAAGTTGCTGTTTCTGAGGCGCGAGCGGCC 1383

892 CCGCGCTGTATCACAGCGCGCGCGCGCGCGCTTGTGTGTGAGCGCGCGGTG 951

1384 ACACCAAGTGCAAAGCCTTGGGACCGCGCGCGCTTATCCACTGGGTAGCCCCCGAT 1443

952 AGCTTGGCTGCGAGCGGTGGTGAACCGCGCGGTGCTGCTGCTGGTGGCACCTGAT 1011

1444 GACCGCTGTGTAGGAACTCTCTCAAGGACCGCTGTCTATGACAATGGGACCTTGGACATC 1503

1012 GGGCGGCTGTGGGAACTTCCAGCGGACCGCGCGCGCGCGCGCGCTTGTGTGATGTG 1071

1504 TTCATCACCATCTCAGGACAGTGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCGGA 1563

1072 ACCATCACCATCTGAGGACAGTGGACCTTCACTTGTATGCTGCTTCAATGCTGCTGG 1131

1564 GAGGCGCGCGCATGTGTGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCAC 1623

1132 GAAGCGCGCGCGCGTGTGAGGTGTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1185

1624 AGCCGCACTGCAACCCCGCAAGTCCCGCTTCTCAGACATCATCTGGGTTCAGCAAGAC 1683



Db	1186	CGCCGGGTGCCCCCGCGCTCTACCGAGCCCGGCTCTCTGACATCGCCACGCCGGC	1245
Qy	1684	CGGGAGGTGGAGGAGTGGGGCGGAGAGCCCTCCAAAAGCCCCCGGAAACGGGCTGTG	1743
Db	1246	AGACAGGTGCCAAGTCTTCCGGTGTAGCGTC-----GGCTC	1284
Qy	1744	CTTGCTGTGAAGTACACACACCTCGGCCCTTGGTCAAGTGTCTGTGAGCAAGTCAGCA	1803
Db	1285	GTGGCAGCCGAGCTCACCTCGAATCCGCTCATCCGCTGGCCAGCCAGAGGCTGTG	1344
Qy	1804	CCCCGGTCAAGATGTACAGCTGACAGTCACTGCTCTGACGATGAGGTACTGATTTAC	1863
Db	1345	CCCGAATACGATGTACAGGTTCACTACACAGTTCGTTGATGACTCCCTCGTCTAC	1404
Qy	1864	AGGATGATCCAGCTCCAAAGGCTTCTGCTGTGTCACCACTGTGTGTCAGGACTGGC	1923
Db	1405	AGGATGATCCCGTCCACCACTGACAGCTTCTCTGTTGAATGACTGTGGCGGGCGCTGC	1464
Qy	1924	TAGCACTTGTGTGCTGGCCATGTGGGATGACAGCCACGACACTCAACGCCACCAAC	1983
Db	1465	TAGCACTTGTGTGCTGGCGGTCTACGACGAGCGGGGCCACAGCGCTGCCGCAACGGA	1524
Qy	1984	ATCGTGGGTGCGCCCGAGTCTTCCACCAAGGCTGACTACCCGCGAGTGCCTCATGCAC	2043
Db	1525	GTGTTGGCTGTGTACAGTTACACACCGCTGGGATCCGGGCGCTGCGCCGCTGAGG	1584
Qy	2044	AGCAGATTCTGGGGGACCATGATCTCTGGTTCATCGGGGGCATCATCTGGCCAGCTG	2103
Db	1585	GCCCATTTCTTGGCGGCGACCATGATCGCCATCGGGGGCGTCTATCGTGGCTCGGTC	1644
Qy	2104	CTGCTTTCATCTCATCTCTCATGCTGCGCTACAAGTCTGCAACACAGGCGCCCGAG-	2162
Db	1645	CTGCTTTCATCTCTGCTCATGATCCGCTATAGGTGTATCGCAGCGGGACGCCGC	1704
Qy	2163	-----CRAGATGGCAGCGGCGCTGAGCAATGTACTGCGAGACCAACGGC	2208
Db	1705	CGGTCAAGGCTCCAGTCTGCTCCCGCGGTGAGCCAGCTGTGCTCGCAGACCAACGGC	1764
Qy	2209	GCCAGCCAGCTTCAAGCAGGCGACAGCGGGCGGCGGCGGCGGCGGCGGCGGCGAAG	2268
Db	1765	GCAGGCACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG	1824
Qy	2269	GTGCTGTGCGCAACAGCTCTTGACTTACCGCCAGCTGCGCGCGCGCGCGCTGACTCC	2328
Db	1825	CGCAGGTGAGTCCAGGCTGCGCCCGCGCTGCGCGTGCAGGCGCAAGGCGGCGG	1884
RESULT 14			
AAD28124			
ID	AAD28124 standard; DNA; 1653 BP.		
XX	AAD28124;		
AC	AAD28124;		
XX	22-APR-2002 (first entry)		
DT	Human leucine-rich repeat-8 (ZLR8) DNA #2.		
DE	Human; leucine-rich repeat-8; ZLR8; cytosstatic; gene therapy; leukaemia;		
XX	endometrial adenocarcinoma; renal cell; colon; prostate; retinoblastoma;		
KW	cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;		
KW	mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;		
KW	rhabdomyosarcoma; genitourinary tract; chromosome 11q13; ds.		
XX	Homo sapiens.		
OS	Key		
XX	FH		
FT	CDS		
FT	Location/Qualifiers		
FT	1..1653		
FT	/*tag= a		
FT	/product= "ZLR8 protein"		
FT	/note= "CDS does not include stop codon"		
FT	/partial		
FT	1..54		
FT	sig_peptide		
FT	/*tag= b		

FT	mat_peptide	55..1653	
FT	/*tag= c		
FT	/product= "Mature ZLR8 protein"		
XX	WO200202604-A2.		
XX	10-JAN-2002.		
XX	02-JUL-2001; 2001WO-US20999.		
XX	30-JUN-2000; 2000US-215446P.		
XX	(ZYMO ) ZYMOGENETICS INC.		
XX	Thayer EC, Sheppard PO, Presnell SR;		
XX	WPI; 2002-154725/20.		
DR	P-PSDB; AAE17484.		
XX	New leucine-rich repeat proteins and polynucleotides, useful for		
XX	diagnosing and treating disorders related to abnormal cell growth e.g.		
XX	retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia,		
XX	kidney and lung tumors		
XX	Claim 19; Page 64-67; 82pp; English.		
XX	The invention relates to an isolated polypeptide comprising leucine-rich		
XX	repeat proteins Zlr7, Zlr8, Zlr9 and spliced variants of Zlr8, Zlr9.		
XX	Zlr7, Zlr8, and Zlr9 proteins are useful in directing the secretion of		
XX	proteins of interest from a host cell and to monitor the secretion of		
XX	proteins in general from cells and tissues. The Zlr8 DNA and proteins		
XX	are useful in diagnosing and treating disorders related to abnormal cell		
XX	growth, including retinoblastoma, renal cell adenocarcinoma, endometrial		
XX	adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphoma, leukemia,		
XX	kidney tumors, germ cell tumors, lung large cell carcinoma, mammary,		
XX	colon adenocarcinoma, genitourinary tract transitional cell tumors,		
XX	rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas		
XX	and prostate adenocarcinoma. Zlr8 protein is useful for identifying		
XX	agonists and antagonists of the polypeptide, for drug design, to screen		
XX	for cell metabolism effecting receptors, for analysis of cell phenotype,		
XX	and as animal feed supplement and cell culture components. Zlr8 DNA is		
XX	also useful in gene therapy. The present sequence is human Zlr8 DNA.		
XX	Zlr8 gene is located on chromosome 11q13.		
XX	Sequence 1653 BP; 223 A; 617 C; 526 G; 287 T; 0 other;		
SQ	Query Match		
	Best Local Similarity, 63.7%; Score 608; DB 24; Length 1653;		
	Matches 1025; Conservative 0; Mismatches 540; Indels 43; Gaps 5;		
Qy	520	GACGCTGCCCCCAAGTACTGTCTGCGCAGAACTCTGTCTGATCTACTGGGACCTGTGC	579
Db	43	GCGGCTGCGCTGCGCTGCGCTGCGCAGAACCTGTCTGAGTCTGCTCAGACCTCTGT	102
Qy	580	CCCTCAAGGGGTGCTCTTGTACCCCTGATATTGACGGCGGACAGTGGAGTGGCG	639
Db	103	GCCACCGAGGCTGTCTTGTGCGGCCCAACGTTGGACCGGCGGACAGTGGAGTGGCG	162
Qy	640	CTGGGCGGCAACTTTCATCATCCACATCAGCCCGCAGGACTTTGCCAAATGACGGGCTG	699
Db	163	CTGCTGAGAACTTTCATCAGGCGCTGGGCGCCCTGACTTCGCGAACATGACGGACTG	222
Qy	700	GTGGACTGACCTGTCTCAGGAACACCATCAGCCATCCAGCCCTTTTCTTCTTGGAC	759
Db	223	GTGGACTGACACTGTCTCGCAATGCCATCAGCCGATGGGGCGGCGCTTTGGGAC	282
Qy	760	CTCAGAGCTCGCTCCCTGCTCTTGTACAGCAATCGGCTGCCAAGCTTGGGAGGAC	819
Db	283	CTCAGAGCTCGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT	342
Qy	820	ACCTTCCGGGGCTGTGTCACCTGACAGCACTTATCGTGAACAAACACAGCTGGGGCGC	879
Db	343	AGCCTCCGGGGCTGTGTCACCTGACAGCACTTATCGTGAACAAACACAGCTGGGGCGC	402



QY 880 ATGCGAGATGAGCGCTTTTGGAGACTTCTGCTGACATTTGGAGGATCTGGACCTCTCTAC 939  
Dd |||||  
403 ATGCGCGCGGAGCGCTTGCAGACTTCTTAGAGAGCTTGAGGACCTTGAGACCTGTCTAC 462  
QY 940 AACAACTCCATGCGCTGCGGTGGGACTTCCGTGCGAGGATGGTCAACCTCCACAGCTG 999  
Dd |||||  
463 AACAACTCCGCGAGGTGCGCTGGCGCGCATCGCGCCATCGCTGCGCTGCACACCTTC 522  
QY 1000 AGCTGGACCAACCTGCTGATCACATCGCGAGGGGACCTTTGCGAGACTGCAGAA 1059  
Dd |||||  
523 AACCTGGACCAATACCTTATTGACGACTGCCCGAGGCGCTTGCAGGCTCGGTGAG 582  
QY 1060 CTGGCGCGCTCGATCTCACTCCAAATCGGCTGCAGAAAGCTGCCCGCTGATCCATCTTT 1119  
Dd |||||  
583 CTCTCCGCTGGACTCACTCCAAACCGCTGGCGCAGCTGGCTGGCGGCGCTTTC 642  
QY 1120 GCCCGCTCCAGGCTCGGCTTTGACAGCCACACCTTTGCCCGCCACCTTTGCTTTAGT 1179  
Dd |||||  
643 TCTGCTGGCGGTAT-----GCAGAGGCGCTCTCCCGCCCGCTTGTGTGAGC 690  
QY 1180 TTTGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTGCTGCGGAGGCTCGAGCG 1239  
Dd |||||  
691 TTTAGCGGAACCCCTGCACTGCAACTGTGAGCTGCTGTGCTGGCGGCTGGCGCG 750  
QY 1240 GACGATGACCTGGAACCTGTGCTGCCAGGGGCGCTCAAGGGTGTCTACTTCTGGCAT 1299  
Dd |||||  
751 CCGGACGACCTGGAACGCTGCCCTCCCGCGCGCTGGCGCGCTACTTCTGGCA 810  
QY 1300 GTGCGTAGGAGAGTTGTGTGCGAGCGCGCTCTCATCACCAGCAGCACACAGTTG 1359  
Dd |||||  
811 GTGCGGAGGCGGAGTTCTCTGTGAGCGCGCTCTCATTTGCCCGCCACGACGCGCTC 870  
QY 1360 CTGGTTCTGAGGGCGGCGGCGCACACTCAAGTCAAAAGCATTGGGGACCCAGCCCG 1419  
Dd |||||  
871 TGGTGTCTGGAAGCGGAGCGGCGGCGCTGCGGTGCGCGGCGCTGGGTGACCGCGCT 930  
QY 1420 CTTATCCACTGGGTAGCCCGCATGACCGCTGGTAGGAACTCTCTCAAGGACCGGTGTC 1479  
Dd |||||  
931 ACCATGCACTGGTGGTCTGTGACGACCGGTTGGTTGGCACTCTCTCCGAGCGCGGT 990  
QY 1480 TATGACATGGACCTCGACATCTTCATCACCACATCTCAGGACAGTGGCTTCACC 1539  
Dd |||||  
991 TTTCCCAACGGACCTTAGAGATTGGGGTGTGACCGCGCTGGGAGCGCTGGGGGCTACAC 1050  
QY 1540 TGCTTGTGCAATGCTGCCGAGAGGCGGCGCATGTGTGAGGTCTCCATGTCGAC 1599  
Dd |||||  
1051 TGCTGCGCACCAACCTGCTGTGAGGCGCACGCGGAGTGAAGTGGGGTGTGGCC 1110  
QY 1600 CTGCCACACTCAGCAACAGCAGCGCACTGCAACCCCGCAAGTCCCGCTCTCAGAC 1659  
Dd |||||  
1111 TTGCCCCATGTTGGGAACAGAGTGCAGGGGGGCGCC-----CGGGCGCTCGGAC 1164  
QY 1660 ATCACTGGCTCCAGCAAGACCGCGGAGGTGGAGGAGTGGGGGCGGAGCGCTCCC 1719  
Dd |||||  
1165 ATCGCGCGCTCGCTCGCACTGTGCGAGGGGTGAGGGAGCGCTGAGTCTGAGCC- 1220  
QY 1720 AAAAGCCCCCGGAGCGGCTGTGCTGTGCTGAAAGTGCACACACCTCGCGCTGTC 1779  
Dd |||||  
1221 -----AGCCGTGAGGTTGACGGAGGTGACCGCGCTGACCGCTCAGGGCTGGTG 1263  
QY 1780 AAGTGTGCTGTGACGAAGTCAAGCCCGGGGTGAAGTGTACAGCTGCAGTACAACTGC 1839  
Dd |||||  
1264 AGCTGGGGTCCCGGGCGGCGGAGCGGAGTGTGAGTGTCCAAATCCAGTACACAGC 1323  
QY 1840 TGTGACATGAGTACTGATTTACAGGATGATCCAGGCTCCAAAGGCGCTTGTGGTGC 1899  
Dd |||||  
1324 AGCGAAGATGAGACCTCATCTACCGGATTGTCCAGCGCTTCCAGCCACCACTTCTGCTG 1383  
QY 1900 AACAACTCGTGTGAGGAGCTGGCTACGACTGTGTGCTGCGCCATGTGGGATGACACA 1959  
Dd |||||  
1384 AAGCACTCGTCCCGCGCGCTGACTATGACCTTGCTGCTGTGCGCTTGTACCGCGCGCT 1443

QY 1960 G---CCAGCAGACTCAGCGCCACCAACATCTGTGGGCTCGGCCAGGTCTTTCACCAAGCT 2016  
Dd |||||  
1444 GGGCGCTCTGACCTCAGCGCCACCAAGGCTGTGGGCTGTGGCCATTTCTCCACGCTGCCG 1503  
QY 2017 GACTACCCGCGAGTGCAGTCCATGCACAGCCAGATTCTGGGCGGACCATGATCTGTGTC 2076  
Dd |||||  
1504 GCGTCCGCGCTGTGCAGCGCTGAGGCCACGCTGTGGGCGGACCTGACCGTGGCC 1563  
QY 2077 ATCGGGGGCATCATCTGGCCACGC-TGCTGGTCTTTCATCGTCACTCT 2123  
Dd |||||  
1564 GTGGGGGTGTGCTGTGCTGCTTACTGTGCTTCACTGTGGCCCT 1611

RESULT 15  
AAD28125  
ID AAD28125 standard; DNA; 1653 BP.  
XX  
AC AAD28125;  
XX  
DT 22-APR-2002 (first entry)  
XX  
Human leucine-rich repeat-8 (ZLR8) degenerate DNA #2.  
DE  
XX  
Human; leucine-rich repeat-8; ZLR8; cytosstatic; gene therapy; leukaemia;  
KW endometrial adenocarcinoma; renal cell; colon; prostate; retinoblastoma;  
KW cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;  
KW mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;  
KW rhabdomyosarcoma; genitourinary tract; chromosome 11q13; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200202604-A2.  
XX  
PD 10-JAN-2002.  
XX  
PF 02-JUL-2001; 2001WO-US20999.  
XX  
PR 30-JUN-2000; 2000US-215446P.  
XX  
PA (ZYMO) ZYMOGENETICS INC.  
XX  
PI Thayer EC, Sheppard PO, Presnell SR;  
XX  
DR WPI; 2002-154725/20.  
XX  
PT New leucine-rich repeat proteins and polynucleotides, useful for  
diagnosing and treating disorders related to abnormal cell growth e.g.  
PT retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia,  
PT kidney and lung tumors  
XX  
PS Disclosure; Page 69-70; 82pp; English.  
XX  
CC The invention relates to an isolated polypeptide comprising leucine-rich  
repeat proteins Zlr7, Zlr8, Zlr9 and spliced variants of Zlr7, Zlr8, Zlr9.  
CC Zlr7, Zlr8, and Zlr9 proteins are useful in directing the secretion of  
CC proteins of interest from a host cell and to monitor the secretion of  
CC proteins in general from cells and tissues. The Zlr7 DNA and proteins  
CC are useful in diagnosing and treating disorders related to abnormal cell  
CC growth, including retinoblastoma, renal cell adenocarcinoma, endometrial  
CC adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphocytic leukaemia,  
CC kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,  
CC colon adenocarcinoma, genitourinary tract transitional cell tumours,  
CC rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas  
CC and prostate adenocarcinoma. Zlr9 protein is useful for identifying  
CC agonists and antagonists of the polypeptide, for drug design, to screen  
CC for cell metabolism effecting receptors, for analysis of cell phenotype,  
CC and as animal feed supplement and cell culture components. Zlr9 DNA is  
CC also useful in gene therapy. The present sequence is human Zlr8  
CC degenerate DNA. Zlr8 gene is located on chromosome 11q13.  
XX  
SQ Sequence 1653 BP; 179 A; 217 C; 329 G; 195 T; 733 other;

Query Match 17.6%; Score 554.6; DB 24; Length 1653;

[illegible]

1043	DB	GNTAYACNTGYATHGCNACNAAYCCNGCGNGGARGACNACNGMNGTNGARYTNMGNG	1102
1592	QY	TCGTCCAGCTGCCACACCTTCAGACAACAGCACCCAGCGGCACTGCACCCCCCAAGTCCCGCC	1651
1103	DB	TNYTNGCNVTCNCAYGCGGNAAYWSNWS-----NGCNGARGGNGMNGCNCNGNC	1156
1652	QY	TCTCAGACATCACTGCTCTCAGCAAGACACCGCGGGGAGGTGGAGCGAGTGGGGCGGAG	1711
1157	DB	CNWSNGAYATHGNCNGCNWSGNCNMGNACNGCNGCNGARGGNGARGGNACTNYTNGARWSNG	1216
1712	QY	AGCCTCCCAAAAGCCCCCGGAACGGGCTGTGCTGTCTCTGAAGTGAACCACTCGG	1771
1217	DB	ARCC-----NGCNGTNCARGTNAACNGCNGACNWSNG	1255
1772	QY	CCCTGGTCAAGTGTCTGTGAGCAACTCAGCACCCCGGGTGAAGATGTACCACTCGACT	1831
1256	DB	GNVYNTGNWSNTGGGNCNGGMMGNCNGCNGAYCCNGTNTGGATGTTTCARATHCART	1315
1832	QY	ACAACTGCTCTGACGATAGGTTACTGATTTCAGGATGATCCAGCCTTCAACAAGGCT	1891
1316	DB	AYAAWSNWSNGARGAYGARACNTYNATHAYMGNATHGNCNGCNSNWSNCACT	1375
1892	QY	TCGTGCTCAACAACCTGGTGTGTCAGGACCTGGCTACGACTTGTGTGCTGGCCATGTGG	1951
1376	DB	TYTYNTYNAARCAYYTNGTNCNGGNCNGAYTAYGAYTNTGYTYNTYNGCNYTNWSNC	1435
1952	QY	ATGACACAGCCACGACA--CTCACGGCCACCAACATCGTGGGCTGGCCGCTTCTTCA	2008
1436	DB	CNGCNGCNGGNCNWSNGAYYTNACNACNMGNTNYTNGGNTGYGCNCAYTYTWSNA	1495
2009	QY	CCAAAGGTGACTACCCCGAGTGCCAGTCCATGCACAGCCAGATCTCTGGGCGGCACCATGA	2068
1496	DB	CNYTNCNGCNSNCCNTYNTGYCAYGCNTNCARGCNCAYGNTNYTNGGNGNACNYTNA	1555
2069	QY	TCCTGCTCATCGGGGGCATCATCGTGCCACGCTGTGCTT	2111
1556	DB	CNGTNGCNGTNGGNGNGTNYTNGTNGCNGNYTNACNGNYT	1598

Search completed: June 8, 2003, 15:46:32  
Job time : 698 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 15:34:52 ; Search time 175 Seconds  
(without alignments)  
5509.667 Million cell updates/sec

Title: US-09-831-846-1  
Perfect score: 3144  
Sequence: 1 gctgtgctccctctcgtga.....gcggttggtgattgattg 3144

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues .

Total number of hits satisfying chosen parameters: 982724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.8	2.6	7452	3 US-08-592-500-1	Sequence 1, Appli
2	81.8	2.6	7452	3 US-08-195-006-1	Sequence 1, Appli
3	81.8	2.6	7452	5 PCT-US94-07644A-1	Sequence 1, Appli
4	81.6	2.6	2852	4 US-09-063-950-1	Sequence 1, Appli
5	79.4	2.5	2019	4 US-09-063-950-3	Sequence 3, Appli
6	77.2	2.5	5176	4 US-09-182-024A-1	Sequence 1, Appli
7	63.8	2.0	4843	3 US-08-986-485-1	Sequence 1, Appli
8	56.4	1.8	2600	2 US-08-427-497E-4	Sequence 4, Appli
9	56.4	1.8	3189	2 US-08-427-497E-3	Sequence 3, Appli
10	56.4	1.8	3774	2 US-08-341-843B-1	Sequence 1, Appli
11	56.4	1.8	3774	2 US-08-427-497E-1	Sequence 1, Appli
12	56.4	1.8	3774	2 US-08-427-497E-2	Sequence 2, Appli
13	56.4	1.8	3888	4 US-08-506-296B-13	Sequence 13, Appli
14	51.6	1.6	4203	2 US-08-866-757-1	Sequence 1, Appli
15	51.6	1.6	4203	4 US-09-153-593-1	Sequence 1, Appli
16	50.2	1.6	1685	6 5340934-1	Patent No. 5340934
17	48.8	1.6	1362	4 US-08-818-112-7	Sequence 7, Appli
18	48.8	1.6	1362	4 US-08-818-111-7	Sequence 7, Appli
19	48.8	1.6	1362	4 US-09-056-556-7	Sequence 7, Appli
20	48.8	1.6	1362	4 US-09-072-596-7	Sequence 7, Appli
21	48.8	1.6	4403765	4 US-09-103-840A-2	Sequence 2, Appli
22	48.8	1.6	4411529	4 US-09-103-840A-1	Sequence 1, Appli
23	48	1.5	503	4 US-09-103-840A-2	Sequence 40, Appli
24	47.6	1.5	3783	4 US-09-297-269-40	Sequence 20, Appli
25	47.4	1.5	906	1 US-08-506-296B-20	Sequence 1, Appli
26	47.4	1.5	906	1 US-07-817-852A-1	Sequence 1, Appli
27	47.4	1.5	1182	4 US-09-385-028-19	Sequence 19, Appli

28	47.4	1.5	11604	4 US-09-385-028-13	Sequence 13, Appli
29	47.4	1.5	15079	4 US-09-385-028-1	Sequence 1, Appli
C 30	47.2	1.5	2338	1 US-08-425-069-1	Sequence 1, Appli
C 31	47.2	1.5	2338	2 US-08-317-844B-1	Sequence 1, Appli
32	46.6	1.5	44377	2 US-08-804-227C-7	Sequence 7, Appli
33	46.6	1.5	44377	2 US-08-804-198-1	Sequence 1, Appli
34	46.4	1.5	8378	5 PCT-US91-09055-1	Sequence 1, Appli
35	46.2	1.5	4134	4 US-09-162-021B-1	Sequence 1, Appli
36	46.2	1.5	4134	4 US-09-687-477-17	Sequence 17, Appli
37	45.8	1.5	2290	4 US-09-131-648-4	Sequence 4, Appli
38	45.4	1.4	891	4 US-09-452-239-1	Sequence 1, Appli
C 39	45.4	1.4	4403765	4 US-09-103-840A-2	Sequence 2, Appli
40	45.2	1.4	1018	4 US-09-452-239-35	Sequence 35, Appli
41	44.8	1.4	1119	4 US-08-987-943-2	Sequence 2, Appli
42	44.8	1.4	1755	4 US-08-987-943-1	Sequence 1, Appli
C 43	44.8	1.4	1926	4 US-09-249-585A-2	Sequence 2, Appli
C 44	44.8	1.4	2580	3 US-09-050-863-2	Sequence 2, Appli
C 45	44.8	1.4	2580	4 US-09-359-081-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-592-500-1  
; Sequence 1, Application US/08592500  
; Patent No. 6005089  
; GENERAL INFORMATION:  
; APPLICANT: Lanza, Francois  
; APPLICANT: Phillips, David R.  
; APPLICANT: Cazenave, Jean-Pierre  
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,500  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/089,455  
FILING DATE: 09-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 12418-28  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7452 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1462..2419  
FEATURE:  
NAME/KEY: CDS

```

LOCATION: 2422..4101
FEATURE:
NAME/KEY: misc_signal
LOCATION: 68..76
OTHER INFORMATION: /function= "Putitive TPA responsive
OTHER INFORMATION: element"
OTHER INFORMATION: /label= TRE
FEATURE:
NAME/KEY: misc_signal
LOCATION: 471..478
OTHER INFORMATION: /function= "Ets-1 cis-acting
OTHER INFORMATION: sequence"
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: misc_signal
LOCATION: 493..502
OTHER INFORMATION: /function= "Ets-1 cis-acting
OTHER INFORMATION: sequence"
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: repeat region
LOCATION: 593..881
OTHER INFORMATION: /rpt_type= "other"
OTHER INFORMATION: /label= Alu
FEATURE:
NAME/KEY: misc_signal
LOCATION: 897..904
OTHER INFORMATION: /function= "Ets-1 cis-acting
OTHER INFORMATION: sequence"
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1142..1149
OTHER INFORMATION: /function= "Spl binding site"
OTHER INFORMATION: /standard_name= "Spl"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 1178..1184
OTHER INFORMATION: /function= "Ets-1 cis-acting
OTHER INFORMATION: sequences"
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1199..1203
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1263..1269
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1285..1289
OTHER INFORMATION: /function= "GATA-1 binding site"
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1321..1326
OTHER INFORMATION: /function= "GATA-1 binding site"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 1365..1372
OTHER INFORMATION: /function= "Ets-1 cis-acting
OTHER INFORMATION: sequences"
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: repeat region
LOCATION: 6133..6440
OTHER INFORMATION: /rpt_type= "other"
OTHER INFORMATION: /label= Alu
FEATURE:
NAME/KEY: misc_signal
LOCATION: 5610..5615
OTHER INFORMATION: /standard_name= "Polyadenylation
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_signal

```

```

LOCATION: 6966..6971
OTHER INFORMATION: /standard_name= "Polyadenylation
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 7224..7229
OTHER INFORMATION: /standard_name= "Polyadenylation
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 7358..7363
OTHER INFORMATION: /standard_name= "Polyadenylation
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..7452
OTHER INFORMATION: /standard_name= "Nucleotide
OTHER INFORMATION: sequence containing the human GPV gene"
US-08-592-500-1
Query Match 2.6%; Score 81.8; DB 3; Length 7452;
Best Local Similarity 50.4%; Pred. No. 1.1e-08;
Matches 200; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
QY 707 TGACCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTCTCTTCTGACCTCAGAGA 766
Db 3233 TGACTCTGTTCGAGAAACCCGCTGGCAGAGTCCCGGGGTGCTCTTCGGGAGATGGGG 3292
QY 767 GCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCTTTGGGAGAGACACCTCC 826
Db 3293 GCCTGCAGGAGCTGTGGCTGAACCGCACCCAGCTGCGCACCTGCCCGCCGCCCTTCC 3352
QY 827 GGGGCTGTCAACCTGCAGCACCTTATCGTGAACAAACACAGCTGGCGGCGATCGGAG 886
Db 3353 GCACCTGACCGCTCGGCTACTTAGGGTGACTCTGAGCCCGCGGTGAGCGCGTTC 3412
QY 887 ATGAGGCTTTGAGGACTTCTCTGTGACATTTGGAGGATCTGACCTCTCTTACAAACAC 946
Db 3413 CGCAGGCGGCTTCCAGGGCTTGGCGAGCTCCAGGTGCTCGCCCTGACCTCCACAGGCC 3472
QY 947 TCCATGGCTGCGGTGGGACTCCGTGCGACGATGTCAGCTCCACCTCCAGCTGAGCTGG 1006
Db 3473 TGACCGCTTCCCGACGGCTTGCTGCGGCGCTCGCAAGCTGCGCAGGTGTCCTGTC 3532
QY 1007 ACCACAACTGCTGGATCAGCTCCGCGGAGGAGGACCTTTGCAGACCTGCAGAACTGGCC 1066
Db 3533 GCGCAACAGGCTGCGGCGCTGCGCCCTGCGCCCTCTTCGCAATCTCAGCAGCTGGAGA 3592
QY 1067 GCCTGATCTCACCTCCATCGGCTGCAGAACTGCC 1103
Db 3593 GCGTCCAGCTCAGACCAACACAGCTGGAGACCTGCC 3629
RESULT 2
US-08-195-006-1
; Sequence 1, Application US/08195006
; Patent No. 6083688
; GENERAL INFORMATION
; APPLICANT: Lanza, Francois
; APPLICANT: Phillips, David R.
; APPLICANT: Cazenave, Jean-Pierre
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/195,006  
FILING DATE: 10-FEB-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/089,455  
FILING DATE: 09-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 12418-28  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7452 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1462..2419  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2422..4101  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 68..76  
OTHER INFORMATION: /function= "Putitive TPA responsive"  
OTHER INFORMATION: element"  
OTHER INFORMATION: /label= TRE  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 471..478  
OTHER INFORMATION: /function= "Ets-1 cis-acting"  
OTHER INFORMATION: sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 493..502  
OTHER INFORMATION: /function= "Ets-1 cis-acting"  
OTHER INFORMATION: sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: repeat region  
LOCATION: 593..881  
OTHER INFORMATION: /rpt\_type= "other"  
OTHER INFORMATION: /label= Alu  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 897..904  
OTHER INFORMATION: /function= "Ets-1 cis-acting"  
OTHER INFORMATION: sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: misc binding  
LOCATION: 1142..1149  
OTHER INFORMATION: /function= "Spl binding site"  
OTHER INFORMATION: /standard\_name= "Spl"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 1178..1184  
OTHER INFORMATION: /function= "Ets-1 cis-acting"  
OTHER INFORMATION: sequences"  
OTHER INFORMATION: /label= Ets-1  
NAME/KEY: TATA\_signal

LOCATION: 1199..1203  
FEATURE:  
NAME/KEY: TATA\_signal  
LOCATION: 1263..1269  
FEATURE:  
NAME/KEY: misc\_binding  
LOCATION: 1285..1289  
OTHER INFORMATION: /function= "GATA-1 binding site"  
FEATURE:  
NAME/KEY: misc binding  
LOCATION: 1321..1326  
OTHER INFORMATION: /function= "GATA-1 binding site"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 1365..1372  
OTHER INFORMATION: /function= "Ets-1 cis-acting"  
OTHER INFORMATION: sequences"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: repeat region  
LOCATION: 6133..6440  
OTHER INFORMATION: /rpt\_type= "other"  
OTHER INFORMATION: /label= Alu  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 5610..5615  
OTHER INFORMATION: /standard\_name= "Polyadenylation"  
OTHER INFORMATION: signal sequence"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 6966..6971  
OTHER INFORMATION: /standard\_name= "Polyadenylation"  
OTHER INFORMATION: signal sequence"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 7224..7229  
OTHER INFORMATION: /standard\_name= "Polyadenylation"  
OTHER INFORMATION: signal sequence"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 7358..7363  
OTHER INFORMATION: /standard\_name= "Polyadenylation"  
OTHER INFORMATION: signal sequence"  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..7452  
OTHER INFORMATION: /standard\_name= "Nucleotide"  
OTHER INFORMATION: sequence containing the human GPV gene"  
US-08-195-006-1  
Query Match 2.6%; Score 81.8; DB 3; Length 7452;  
Best Local Similarity 50.4%; Pred. No. 1.1e-08;  
Matches 200; Conservative 0; Mismatches 197; Indels 0; Gaps 0;  
QY 707 TGACCCCTGTCAGGAACACCATCGCCACATCCAGCCCTTTTCTTTCTGGACCTCGAGA 766  
|||  
Db 3233 TGACTCTGTTTCAGAAACCCGCTGGCAGAGCTCCGGGGTGCTTTCGGGGAGATGGGG 3292  
QY 767 GCCTCGCTCCCTGCATCTTGACACCAATCGGCTCCAGAGCTTGGGAGGACACCTCC 826  
|||  
Db 3293 GCCTCAGGAGCTGTGGCTGAACCCAGCTCGCACCCCTGCCCGCGGCTTCC 3352  
QY 827 GGGGCTGGTCAACCTGCAGCACCTTATCGTCAACAACACAGCTGGCGGCGATCGCAG 886  
|||  
Db 3353 GCAACCTGAGCCGCTTCGGTACTTAGGGGTGACTCTGAGCCCGCGGCTGAGCGGCTTC 3412  
QY 887 ATGAGGCTTTTGAGGACTTCTCTGTGACATTTGAGGATCTGACCTTCTCTACAAACC 946  
|||  
Db 3413 CGCAGGCGCCTTCCAGGGCCTTGGCAGCTCCAGGTGCTCGCCCTGCACCTCAACGGCC 3472  
QY 947 TCCATGGCTGCCGTGGGACTCCGTGGCAGCGATGGTCAACCTCCACAGCTGAGCTGG 1006  
|||  
Db 3473 TGACCGCCCTCCCGACGCGCTTGTGCGCGGCTCGGCAAGCTGGCGCAGGTGTCCCTGC 3532

```
QY      1007 ACCACAACTGTCGTGCATCATCGCCGGAGGCCACTTTTGCAGACCTGCAGAAAACCTGGCCC   1066  
         |||||    |||||       |     |||||  
Db      3533 GCCGCAACAGGTGGCGCCTGGCCCCCTTGCCGAATCTCAGCACCTGGAGA   3592  
  
QY      1067 GCTCGATCTCACCTCCAATCGCTGCAGAACCTGCC   1103  
         |||||    |||||       |     |||||  
Db      3593 GCCTCAGCTGCACCACACACACCTGGAGACCTGCC   3629  
  
RESULT 3  
PCT-US94-07644A-1  
Sequence 1, Application PC/TUS9407644A  
GENERAL INFORMATION:  
APPLICANT: COR Therapeutics, Inc.  
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Khourie and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07644A  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 012418-003000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7452 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1462..2419  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2422..4101  
FEATURE:  
NAME/KEY: misc_signal  
LOCATION: 68..76  
OTHER INFORMATION: /function= "Putative TPA responsive element"  
OTHER INFORMATION: /label= TRE  
FEATURE:  
NAME/KEY: misc_signal  
LOCATION: 471..478  
OTHER INFORMATION: /function= "Ets-1 cis-acting sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: misc_signal  
LOCATION: 493..502  
OTHER INFORMATION: /function= "Ets-1 cis-acting sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:
```

---

```
/ NAME/KEY: repeat_region  
/ LOCATION: 593..881  
/ OTHER INFORMATION: /rpt_type= "other"  
/ OTHER INFORMATION: /label= Alu  
/ FEATURE:  
/ NAME/KEY: misc_signal  
/ LOCATION: 897..904  
/ OTHER INFORMATION: /function= "Ets-1 cis-acting sequence"  
/ OTHER INFORMATION: /label= Ets-1  
/ FEATURE:  
/ NAME/KEY: misc_binding  
/ LOCATION: 1142..1149  
/ OTHER INFORMATION: /function= "Spl binding site"  
/ OTHER INFORMATION: /standard_name= "Spl"  
/ FEATURE:  
/ NAME/KEY: misc_signal  
/ LOCATION: 1178..1184  
/ OTHER INFORMATION: /function= "Ets-1 cis-acting sequences"  
/ OTHER INFORMATION: /label= Ets-1  
/ FEATURE:  
/ NAME/KEY: TATA_signal  
/ LOCATION: 1199..1203  
/ FEATURE:  
/ NAME/KEY: TATA_signal  
/ LOCATION: 1263..1269  
/ FEATURE:  
/ NAME/KEY: misc_binding  
/ LOCATION: 1285..1289  
/ OTHER INFORMATION: /function= "GATA-1 binding site"  
/ FEATURE:  
/ NAME/KEY: misc_binding  
/ LOCATION: 1321..1326  
/ OTHER INFORMATION: /function= "GATA-1 binding site"  
/ FEATURE:  
/ NAME/KEY: misc_signal  
/ LOCATION: 1365..1372  
/ OTHER INFORMATION: /function= "Ets-1 cis-acting sequences"  
/ OTHER INFORMATION: /label= Ets-1  
/ FEATURE:  
/ NAME/KEY: repeat_region  
/ LOCATION: 6133..6440  
/ OTHER INFORMATION: /rpt_type= "other"  
/ OTHER INFORMATION: /label= Alu  
/ FEATURE:  
/ NAME/KEY: misc_signal  
/ LOCATION: 5610..5615  
/ OTHER INFORMATION: /standard name= "polyadenylation signal sequence"  
/ FEATURE:  
/ NAME/KEY: misc_signal  
/ LOCATION: 6966..6971  
/ OTHER INFORMATION: /standard name= "Polyadenylation signal sequence"  
/ OTHER INFORMATION: /label= Ets-1  
/ FEATURE:  
/ NAME/KEY: misc_signal  
/ LOCATION: 7224..7229  
/ OTHER INFORMATION: /standard name= "Polyadenylation signal sequence"  
/ OTHER INFORMATION: /label= Ets-1  
/ FEATURE:  
/ NAME/KEY: misc_signal  
/ LOCATION: 7358..7363  
/ OTHER INFORMATION: /standard name= "Polyadenylation signal sequence"  
/ OTHER INFORMATION: /label= Ets-1  
/ FEATURE:  
/ NAME/KEY: misc_feature  
/ LOCATION: 1..7452  
/ OTHER INFORMATION: /standard name= "Nucleotide sequence containing the human GPV gene"  
/ OTHER INFORMATION: /label= Ets-1
```

PCT-US94-07644A-1

Query Match 2.6%; Score 81.8; DB 5; Length 7452;  
Best Local Similarity 50.4%; Pred. No. 1.1e-08;  
Matches 200; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 707 TGACCTGTCCAGGAACACCATCAGCACATCCAGCCCTTTTCTTGGACCTCGAGA 766  
DB 3233 TGACTCTGTTCAGAAACCCGCTGGCAGAGCTCCGGGGGTGCTCTTCGGGAGATGGGG 3292

QY 767 GCCTCGCTCCTGTCATTTGACGAATCGCTGCCAAGCCTTGGGAGGACACCTCC 826  
DB 3293 GCCTGAGGAGCTGTGGCTGAACCGCACCTGCGCACCTGCCCGCGCGCTTCC 3352

QY 827 GGGGCTGTCAACCTGAGACCTTATCTGAACAAACAGCTGGCGGCATCGAG 886  
DB 3353 GCAACCTGAGCGCCCTGCGGTACTTAGGGGTGACTCTGAGCCCGCGCTGAGCGGCTTC 3412

QY 887 ATGAGCTTTTGAGGACTTCTGCTGACATTTGAGGATCTGACACTCTCTACAACACC 946  
DB 3413 CCGAGGCGCCTTCAGGGCCTTGGGAGCTCAGGTGCTCGCCTGCACTCAAAGGCC 3472

QY 947 TCCATGGCTGCGGTGGGACTCCGTGCGAGCATGTCAACCTCCACAGCTGAGCCTGG 1006  
DB 3473 TGACGCCCTCCCGAGGCTTGTGCGGCGCTCGGCAAGCTGCGCAGGTGTCCCTGC 3532

QY 1007 ACCACAACTGCTGATCATCGCGGAGGACCTTTGCAACCTGCAAGAACTGGCCC 1066  
DB 3533 GCGCAACAGGCTGCGCGCCCTGCCCCGTGCTCTTCGCAATCTCAGCAGCCTGGAGA 3592

QY 1067 GCTGTGATCTCACTCCATCGCTGACAGCTGCC 1103  
DB 3593 GGTCTCAGCTGACCAACACAGCTGGAGACCTGCC 3629

RESULT 4

US-09-063-950-1  
; Sequence 1, Application US/09063950C  
; Patent No. 6225085  
; GENERAL INFORMATION:  
; APPLICANT: Holtzman, Douglas A.  
; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: MEI-019  
; CURRENT APPLICATION NUMBER: US/09/063,950C  
; CURRENT FILING DATE: 1998-04-21  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2852  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (160)..(2178)  
US-09-063-950-1

Query Match 2.6%; Score 81.6; DB 4; Length 2852;  
Best Local Similarity 46.6%; Pred. No. 8.9e-09;  
Matches 338; Conservative 0; Mismatches 379; Indels 9; Gaps 2;

QY 315 CCGGAGCCGCGGAGGGTTAGTCTGAGTCTGTCTCGGCGGGGGAAGGATGCTGGCCGA 374  
DB 15 CCGGAGCCGCGGCGGGTGGACGCGACTCGAACGAGTTGCTTCGGGACCCAGGACCCC 74

QY 375 GCGGGGAGCCGCGGCGCCCCCGGAGCGGCTCGGTGCCACCCAGCGGGGTAGATG 434

DB 75 CTGGGCGCGACCCCGCAGAAAGACTGAGGCGCGGCTGCCCGCCCGGCTCCCTGCG 134

QY 435 CTGCTCTCGCCAGCGCTGAGTGACCAAGCATGAGAGACCTGCTTTGGTGGCTGTAGC 494

DB 135 CCGCGCGCGCTCCCG- - - - -GGACAGAAGATGCTCTCAGGGTCCCTCTGCTGCTGC 188

QY 495 GTTTGCGATGGCGTTTGGCGTGTGTCAGCGCTGCCCAAGTACTGTGTCTGCGAGATCT 554

DB 189 GCTGCTCCTCTACTGCGCCCTGGGGCTGGGGTGCAGGGCTGCCCATCGGCTGCCAGTG 248

QY 555 GTCTCAGTCACTGGGGACCTGTGCCCCCTCAAGGGGTGCTCTTTGTACCCCTGATAT 614

DB 249 CAGCGAGCCACAGACAGTCTTCTGCACTGCCCGCAGGGGACACCGTGGCCGAGACGT 308

QY 615 TGACCGGGGACAGTGGAGCTGCGCTGGGCGGCACTTCACTATCCACATCAGCGGCA 674

DB 309 GCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGCGAG 368

QY 675 GGACTTTCCCAACATGACGGGCTGGTGGACCTGACCTGCTCCAGGAACACCATCAGCCA 734

DB 369 CAGCTTTGCGGCGCTGCGGGCTGCAGCTCTCTGACCTGTGACAGAACAGATCGCCAG 428

QY 735 CATCAGCGCCTTTTCTTCTTGACCTCGAGAGCCTCCGCTCCCTGCACTTTGACAGCAA 794

DB 429 CCGTCCCGAGCGGGTCTTCCAGCCACTCGCAACCTCAGCAACCTGGACCTGACGGCCAA 488

QY 795 TCGGCTGCCAAGCCTTGGGGAGGACACCTCGGGGCTGTGTAACCTGCGAGCCTTAT 854

DB 489 CAGGCTGATGAAATCACAATGAGACCTTCGCTGGCGCTGCGGCGCTCGAGCGCTCTA 548

QY 855 CGTGAACACACAGCTGGGCGCATCGAGATGAGGCTTTTGAAGACTTCTCTGCTGAC 914

DB 549 CTTGGCAAGAACCCCATCCGCCACATCCAGCTGTGCTTCGACACGCTCGACCGCT 608

QY 915 ATTGAGGATCTGGACCTCTCTCAACAACCTCCATGGCCTGCGTGGGACTCCGTTGG 974

DB 609 CTTGGAG- - -CTCAAGCTGCAGGACACAGAGCTGGGGCACTGCCCGCTGCGCTGCC 665

QY 975 AGCATGTCAACCTCCACAGCTGAGCTGGACCAACACCTGCTGGATCAGATCGCGA 1034

DB 666 CCGCTGCTGCTGTGACCTCAGCCACAACAGCCTCTCGGCGCTCGAGCGCGCATCCT 725

QY 1035 GGGCAC 1040  
DB 726 GGACAC 731

RESULT 5

US-09-063-950-3  
; Sequence 3, Application US/09063950C  
; Patent No. 6225085  
; GENERAL INFORMATION:  
; APPLICANT: Holtzman, Douglas A.  
; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: MEI-019  
; CURRENT APPLICATION NUMBER: US/09/063,950C  
; CURRENT FILING DATE: 1998-04-21  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2019  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2019)  
US-09-063-950-3

Query Match 2.5%; Score 79.4; DB 4; Length 2019;  
Best Local Similarity 49.3%; Pred. No. 2.4e-08;  
Matches 237; Conservative 0; Mismatches 241; Indels 3; Gaps 1;

QY 560 AGTCACTGGGAGCCCTGTGCCCCCTCAAGGGGCTCTTTGTACCCCTGATATGACC 619

DB 95 AGCCACAGACAGTCTTCTGCACTGCCCGCAGGGGACCAACGCTGCCCGGAGACGTGCCAC 154

QY 620 GCGGACAGTGGAGCTGCGCTGGGCGCAACTTCACTATCCACATCAGCGCCAGGACT 679

DB 155 CCGACACGGTGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGCGACGCT 214



us-09-831-846-1.rni

Db 589 CTTCCAGAAATGCTTTTCCAG---AGCACGCCGGAAGCTCACACAGACTAGATTGAGTGAA 645  
 QY 940 AACAACTCTCATTGGCTGCGTGGGACTCCGTCGACGATGGTCAACCTCCACGAGCTG 999  
 Db 645 AACCATGACGAGGATCCGAGGAAGGCTTCCGCGCATCACCGATGTGAAGAACCTG 705  
 QY 1000 AGCTGGACCAACATGCTGGATCAGATCGCGAGGACCTTTGACAGCTTCGAGAAA 1059  
 Db 706 CAACTGGACAAACACATCAGCTGCATTGAAGATGGAGCTTCCGAGCGCTGCCGAT 765  
 QY 1060 CTGCGCCGCTGGATCTCACCTCCAA 1085  
 Db 766 TTGGAGATCTTACCCTCAACAA 791

RESULT 7

US-08-986-485-1  
 ; Sequence 1, Application US/08986485  
 ; Patent No. 6046030  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WU, SHULIAN  
 ; APPLICANT: SWEET, RAYMOND  
 ; APPLICANT: TRUNEH, ALEMESEGED  
 ; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: RATNER & PRESTIA  
 ; STREET: P.O. BOX 980  
 ; CITY: VALLEY FORGE  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19482  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/986.485  
 ; FILING DATE: 08-DEC-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/059,448  
 ; FILING DATE: 22-SEP-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: PRESTIA, PAUL F  
 ; REGISTRATION NUMBER: 23,031  
 ; REFERENCE/DOCKET NUMBER: GH-70264  
 ; TELEPHONE: 610-407-0700  
 ; TELEFAX: 610-407-0701  
 ; TELEX: 846169  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4843 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; US-08-986-485-1

Query Match 2.0%; Score 63.8; DB 3; Length 4843;  
 Best Local Similarity 49.1%; Pred. No. 7.7e-05;  
 Matches 198; Conservative 0; Mismatches 202; Indels 3; Gaps 1;  
 QY 684 CAACATACGCGGCTGGTGGACCTGCTCCAGAACACCATCAGCCATCCAGCC 743  
 Db 669 CAAGCTACCCAGGCTGACACAACTGGACCTCAATCGAAACAGGATTCGGCTGATAGAGG 728  
 QY 744 CTTTTCCTTTCTGGACCTCGAGAGCTCCGCTCCCTCGCATCTTGACAGCAATCGGTGCC 803  
 Db 729 CCTCACCTCCAGGGGCTCAACAGCTTGGAGGTCTGAAGCTTCAGCGAAACAACATCAG 788

Tue Jun 10 08:57:37 2003

QY 680 TTGCCAATGACGGGGCTGTGGACCTGACCTGCTCCAGGAACACCATCAGCCACATCC 739  
 Db 215 TTGCCGGGCTGCGGGGCTGACGCTCTGGACCTGTCAAGAACAGATCGCCAGCTGC 274  
 QY 740 AGCCCTTTTCTTTCTGGACCTCGAGAGCTCCGCTCCCTGCTGATCTTGACAGCAATCGGC 799  
 Db 275 CCAGCGGGGTCTTCAGCCACTCGCCAACTCAGCAACTGAGACCTGACGCGCAACAGGC 334  
 QY 800 TGCCAAAGCTTGGGAGGACACCTTCGCGGGCTGTCAACTGACGACCTTATCTGGA 859  
 Db 335 TGCATGAATACCAATGAGACTTCCGTCGGCTCGCGGCTCGAGGCTCTTACCTGG 394  
 QY 860 AGAACCAAGCTGGCGGCATCGCAGATGAGCTTTTGAGGACTTCTGCTGATGATGG 919  
 Db 395 GCAAGAACCGCATCCGCGCATCCAGCTGGCTGGCTTGCACACGCTCGACGCGCTCTGG 454  
 QY 920 AGGATCTGGACCTCTCTCAACAACTCATGAGCTGCGCTGGGCTGAGCTCCGTCGAGCA 979  
 Db 455 AG--CTAAGCTGCAGGACACGAGCTGCGGCACTGCGCCCTGCGCTGCGCCCGCC 511  
 QY 980 TGTCAACCTCCAGGCTGAGCTGAGCTGAGCAACCTGCTGGATCAGATCGCGCGAGGCA 1039  
 Db 512 TGCTGCTGGACCTCAGCCACAAAGCTCTGCGCTGGAGCCCGGCATCTCTGGACA 571  
 QY 1040 C 1040  
 Db 572 C 572

RESULT 6

US-09-182-024A-1  
 ; Sequence 1, Application US/09182024A  
 ; Patent No. 6342370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Connolly, Timothy  
 ; APPLICANT: Rajput, Bhanu  
 ; TITLE OF INVENTION: Human Slit Polypeptide and Polynucleotides Encoding  
 ; TITLE OF INVENTION: Same  
 ; FILE REFERENCE: 640100-271  
 ; CURRENT APPLICATION NUMBER: US/09/182.024A  
 ; CURRENT FILING DATE: 1998-10-29  
 ; PRIOR APPLICATION NUMBER: 60/063,946  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/096,420  
 ; PRIOR FILING DATE: 1998-08-13  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 5176  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: cDNA coding for  
 ; OTHER INFORMATION: the human slit polypeptide of SEQ ID NO:2  
 ; US-09-182-024A-1

Query Match 2.5%; Score 77.2; DB 4; Length 5176;  
 Best Local Similarity 51.8%; Pred. No. 9.5e-08;  
 Matches 200; Conservative 0; Mismatches 183; Indels 3; Gaps 1;  
 QY 700 GTGGACCTGACCTGTCCAGGAACACCATCAGCCATCCAGCCCTTTTCTTTCTGGAC 759  
 Db 409 GAGGCTTGAACCTGGACAGAAATATATACCAAGATGACATCCAGCTTCCTGGG 468  
 QY 760 CTCGAGAGCTCCGCTCCCTGCTGATCTTGACAGCAATCGCTGCCAAGCTTCGGGAGGAC 819  
 Db 469 CTCGAAGAACCTCCGAGTCTTGTGATCTGGAAGAACACAGGTACAGCTATCAGAGAGGC 528  
 QY 820 ACCCTCGGGGCTGTGCACTGACGACCTTATCGTGAACAAACACAGCTGGGCGGC 879  
 Db 529 GCCTTCCAGGACCTTGAAGAGCTAGAGGACTCGGCTGAACAAAGATAAGCTGCAAGTC 588  
 QY 880 ATCGCAGATGAGGCTTTTGGAGACTTCTCTGCTGACATGGAGATCTGGACCTCTCTAC 939

```

QY 804 AAGCCTTGGGGAGGACACCTCGGGGCTGGTCAACCTGCGACACCTTATCGTGAACAA 863
Db 789 CAAACTGACAGATGGGGCTCTTGGGACTGTCCAAATGCAATGCTGCTGACCTGGAGTA 848
QY 864 CAACAGCTGGGGGATCGCAGATGAGCTTTTGGAGACTTCTCTGCTACATTTGGAGGA 923
Db 849 CGACAGCTGGTAGAAGTGAACAGCGGCTCGCTACGGCTCACGG---CCTGCAATCA 905
QY 924 TCTGGACCTCTCTACAAACCTCCATGGCTGCGCTGGGACTCGCTGCGACGATGCT 983
Db 906 GTCCACCTCAGCAACAATTCATCGCTCGCATTCACCAAGGGTGGAGCTTGTGCA 965
QY 984 CAACCTCCACGCTGAGCTGACCAACACCTGCTGGATCAATCGCGGAGGACCTT 1043
Db 966 GAAGCTGCAATGAGTGGTCTGCTTCAACACCTGACACGGCTGACAGGAGAGGCT 1025
QY 1044 TCGACACTGCAAGAACTGGCGGCTGGATCTCACTCCAAT 1086
Db 1026 GGCCGAGCTGAGAGCTGAGTGTCTGCTGCTCAAGCCCAAT 1068

```

RESULT 8

```

US-08-427-497E-4
; Sequence 4, Application US/08427497E
; Patent No. 5969124
; GENERAL INFORMATION:
; APPLICANT: Lemmon, Vance
; TITLE OF INVENTION: A Method for Characterizing the
; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
; Patent No. 5969124
; TITLE OF INVENTION: the Nucleotide Sequence
; TITLE OF INVENTION: Characterized Thereby
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fay, Sharpe, Beall, Pagan,
; ADDRESSEE: Minnich & McKee
; STREET: 1100 Superior Avenue
; STREET: Suite 700
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2518
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; MEDIUM TYPE: storable
; COMPUTER: Compaq Prolinea 5100e
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,497E
; FILING DATE: April 24, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/904,991
; FILING DATE: June 26, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Minnich, Richard J.
; REGISTRATION NUMBER: 24,175
; REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 861-5582
; TELEFAX: (216) 241-1666
; TELEX: (216) 980162
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2600
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acids
; HYPOTHETICAL: irrelevant
; ANTI-SENSE: no

```

```

; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: 17-18 week fetus
; IMMEDIATE SOURCE:
; LIBRARY: Stratagene cDNA Library 936206
; CLONE: 4
; PUBLICATION INFORMATION:
; AUTHORS: Hlavin, Mary Louise
; AUTHORS: Lemmon, Vance
; TITLE: Molecular structure and functional testing of
; TITLE: human L1CAM: an interspecies comparison.
; JOURNAL: GENOMICS
; VOLUME: 11
; PAGES: 416-423
; DATE: 1991
; RELEVANT RESIDUES IN SEQ ID NO: 1108 to 3708
US-08-427-497E-4
Query Match 1.8%; Score 56.4; DB 2; Length 2600;
Best Local Similarity 49.3%; Pred. No. 0.0026;
Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
QY 1334 TCATCACCAGCACACACAAAGTTGCTGTTCTGGAGGGCCAGGGCCACACTCAAGT 1393
Db 176 TCCTGACTGCGGACAATCAGACGTATGCTGCTCCAGGGCAGCACTGCTTCTCT 235
QY 1394 GCAAGAGCATTTGGGACCCCGCTTATCCACTGGGTAGCCCGGATGACCGCTCG 1453
Db 236 GCAAGGCTTCGGAGCGGCTGTGCCAGGTTCAGTGGCTGGAGGATGGACAACAG 295
QY 1454 TAGGGAACCTCCTCAAGGACCGCTGTCTATGACAAATGGCACCTGGACATCTT 1513
Db 296 TCGTTTCAGGACGACGCTTCTTCCCTATGCCAATGGGACCTGGGCACTCC 355
QY 1514 CATCTCAGACAGTGGTCTTCACTGCAATTCGCAATGCTGCGGAGAGGCCACGG 1573
Db 356 AGGCCAATGACACCGGAGCTACTTCTGCTGGTCCCAATGACCAAAACAATGTTACCA 415
QY 1574 CCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACTCAGCAACAGCACCGCCGAC 1631
Db 416 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCGAGCAC 473

```

RESULT 9

```

US-08-427-497E-3
; Sequence 3, Application US/08427497E
; Patent No. 5969124
; GENERAL INFORMATION:
; APPLICANT: Lemmon, Vance
; TITLE OF INVENTION: A Method for Characterizing the
; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
; Patent No. 5969124
; TITLE OF INVENTION: the Nucleotide Sequence
; TITLE OF INVENTION: Characterized Thereby
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fay, Sharpe, Beall, Pagan,
; ADDRESSEE: Minnich & McKee
; STREET: 1100 Superior Avenue
; STREET: Suite 700
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2518
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; MEDIUM TYPE: storable
; COMPUTER: Compaq Prolinea 5100e
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,497E

```

Patent No. 5872225 Nucleotide Sequence of L1CAM and the Nucleotide Sequence Characterized Thereby

TITLE OF INVENTION: Patent No. 5872225

TITLE OF INVENTION: the Nucleotide Sequence Characterized Thereby

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS: FAY, SHARPE, BEALL, FAGAN, ADDRESSSEE: MINNICH & MCKEE

STREET: 1100 Superior Avenue

SUITE: 700

CITY: Cleveland

STATE: Ohio

COUNTRY: U.S.A.

ZIP: 44114-2518

COMPUTER READABLE FORM: Diskette, 3.50 inch, 720 Kb

MEDIUM TYPE: storable

COMPUTER: Compaq Prolinea 5100e

OPERATING SYSTEM: DOS 5.0

SOFTWARE: ASCII

CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/341,843B

FILING DATE: No. 5872225ember 18, 1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA: APPLICATION NUMBER: 07/904,991

FILING DATE: June 26, 1992

ATTORNEY/AGENT INFORMATION: NAME: Minnich, Richard J.

REGISTRATION NUMBER: 24,175

REFERENCE/DOCKET NUMBER: CWR 2 149-1

TELECOMMUNICATION INFORMATION: TELEPHONE: (216) 861-5582

TELEFAX: (216) 241-1666

TELEX: (216) 980162

INFORMATION FOR SEQ ID NO: 1: SEQUENCE CHARACTERISTICS: LENGTH: 3774

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: irrelevant

ANTI-SENSE: no

ORIGINAL SOURCE: ORGANISM: Homo sapiens

INDIVIDUAL ISOLATE: 17-18 week fetus

IMMEDIATE SOURCE: LIBRARY: Stratagene cDNA Library 936206

CLONE: synthesis of 4 clones

PUBLICATION INFORMATION: AUTHORS: Hlavin, Mary Louise

AUTHORS: Lemmon, Vance

TITLE: Molecular structure and functional

TITLE: testing of human L1CAM: an

TITLE: interspecies comparison.

JOURNAL: GENOMICS

VOLUME: 11

ISSUE: 416-423

PAGES: 1991

DATE: 1991

RELEVANT RESIDUES IN SEQ ID NO: 1 to 3774

US-08-341-843B-1

Query Match 1.8%; Score 56.4; DB 2; Length 3774;

Best Local Similarity 49.3%; Pred. No. 0.0028;

Mismatches 151; Indels 0; Gaps 0;

Matches 147; Conservative

Qy 1334 TCATCACCAGCACACACAACTTGTCTTGGAGGGCCAGGGCCACACTCAAGT 1393

Db 1283 TCCTGACTGGGACATCAGACCTACATGCTCTCCAGGGCAGCACTGCCTACCTTCTGT 1342

Qy 1394 GCAAAGCCATTGGGAGCCCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCGCTGG 1453

RESULT 10

US-08-341-843B-1

Sequence 1, Application US/08341843B

Patent No. 5872225

GENERAL INFORMATION: APPLICANT: Lemmon, Vance

TITLE OF INVENTION: A Method for Characterizing the

Db 1343 GCAAGCGCTTCGAGCGCTGTGCCAGTGTTCACTGGCTGACGAGGATGGCAACAG 1402  
QY 1454 TAGGGAACCTCTCAAGGACCGCTGTATGACAAATGGCACCTTGACATCTTTCATCACCA 1513  
Db 1403 TCGTTCAGGACGACGCTCTTCCCTGTATGCCAATGGGACCTTGGCATTCAGACCTCC 1462  
QY 1514 CATCTCAGGACAGTGGTGGCTTTCACCTGCAITGCTGCCAATGCTCCGGAGGCGCACGG 1573  
Db 1463 AGGCCAATGACACCGGACGCTACTTCTGCTGGCTGCCAATGACCAAAACAATGTTACCA 1522  
QY 1574 COATGGTGGAGTCTCCATGCTCCAGCTGCCACACTGACCAACAGCAGCAGCCGAC 1631  
Db 1523 TCATGGCTAACTGAAGGTTAAAGATGCAATCAGATCACTCAGGGGCCCCGCGAC 1580

RESULT 11  
US-08-427-497E-1  
; Sequence 1, Application US/08427497E  
; Patent No. 5969124  
; GENERAL INFORMATION:  
; APPLICANT: Lemmon, Vance  
; TITLE OF INVENTION: A Method for Characterizing the  
; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and  
; Patent No. 5969124  
; TITLE OF INVENTION: the Nucleotide Sequence  
; TITLE OF INVENTION: Characterized Thereby  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fay, Sharpe, Beall, Fagan,  
; ADDRESSEE: Minnich & McKee  
; STREET: 1100 Superior Avenue  
; STREET: Suite 700  
; CITY: Cleveland  
; STATE: Ohio  
; COUNTRY: U.S.A.  
; ZIP: 44114-2518

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; MEDIUM TYPE: storable  
; COMPUTER: Compaq Prolinea 5100e  
; OPERATING SYSTEM: DOS 5.0  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: 07/904,991  
; FILING DATE: June 26, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Minnich, Richard J.  
; REGISTRATION NUMBER: 24,175  
; REFERENCE/DOCKET NUMBER: CWR 2 149-3-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (216) 861-5582  
; TELEFAX: (216) 241-1666  
; TELEX: (216) 980162  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3774  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: nucleic acids  
; HYPOTHETICAL: irrelevant  
; ANTI-SENSE: no  
; ORIGINAL SOURCE:  
; ORGANISM: Homo Sapiens  
; INDIVIDUAL ISOLATE: 17-18 week fetus  
; IMMEDIATE SOURCE:  
; LIBRARY: Stratagene cDNA Library 936206  
; CLONE: synthesis of 4 clones

PUBLICATION INFORMATION:  
; AUTHORS: Hlavin, Mary Louise  
; AUTHORS: Lemmon, Vance  
; TITLE: Molecular structure and functional  
; TITLE: testing of human L1CAM: an  
; TITLE: interspecies comparison.  
; JOURNAL: GENOMICS  
; VOLUME: 11  
; ISSUE:  
; PAGES: 416-423  
; DATE: 1991  
; RELEVANT RESIDUES IN SEQ ID NO: 1 to 3774  
US-08-427-497E-1  
Query Match 1.8%; Score 56.4; DB 2; Length 3774;  
Best Local Similarity 49.3%; Pred. No. 0.0029;  
Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;  
QY 1334 TCATCACCAGCAGACACACAAAGTTGCTGTGAGGGGCGAGGGCCACACTCAAGT 1393  
Db 1283 TCCTGACTGCGGACAAATCAGACGTATCGGTGTCCAGGGCAGCACTGCCTACTTCTGT 1342  
QY 1394 GCAAGGCGCATTTGGGACCCCGCCCTTATCCACTGGGTAGCCCGCATGACCGCTGG 1453  
Db 1343 GCAGGCGCTTCGGAGCGCTGTGCGCCAGTGTTCAGTGGCTGACGAGGATGGGACAACAG 1402  
QY 1454 TAGGGAACCTCTCAAGGACCGCTGTATGACAAATGGCACCCCTGGACATCTTTCATCACCA 1513  
Db 1403 TCGTTCAGGACGACGCTCTTCCCTATGCCAATGGGACCTTGGCATTCAGACCTCC 1462  
QY 1514 CATCTCAGGACAGTGGTGGCTTTCACCTGCAITGCTGCCAATGCTCCGGAGGCGCACGG 1573  
Db 1463 AGGCCAATGACACCGGACGCTACTTCTGCTGGCTGCCAATGACCAAAACAATGTTACCA 1522  
QY 1574 COATGGTGGAGTCTCCATGCTCCAGCTGCCACACTGACCAACAGCAGCAGCCGAC 1631  
Db 1523 TCATGGCTAACTGAAGGTTAAAGATGCAATCAGATCACTCAGGGGCCCCGCGAC 1580

RESULT 12  
US-08-427-497E-2  
; Sequence 2, Application US/08427497E  
; Patent No. 5969124  
; GENERAL INFORMATION:  
; APPLICANT: Lemmon, Vance  
; TITLE OF INVENTION: A Method for Characterizing the  
; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and  
; Patent No. 5969124  
; TITLE OF INVENTION: the Nucleotide Sequence  
; TITLE OF INVENTION: Characterized Thereby  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fay, Sharpe, Beall, Fagan,  
; ADDRESSEE: Minnich & McKee  
; STREET: 1100 Superior Avenue  
; STREET: Suite 700  
; CITY: Cleveland  
; STATE: Ohio  
; COUNTRY: U.S.A.  
; ZIP: 44114-2518  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; MEDIUM TYPE: storable  
; COMPUTER: Compaq Prolinea 5100e  
; OPERATING SYSTEM: DOS 5.0  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/427,497E  
; FILING DATE: April 24, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/904,991  
; FILING DATE: June 26, 1992

us-09-831-846-1.rni

Tu Jun 10 08:57:37 2003

ADDRESSEE: The Scripps Research Institute  
STREET: 10550 No. 6313265th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: U.S.  
ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/506,296B  
FILING DATE: 24-JUL-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 488.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 554-2937  
TELEFAX: (619) 554-6312  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3888 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 12..3773  
US-08-506-296B-13

Query Match 1.8%; Score 56.4; DB 4; Length 3888;  
Best Local Similarity 49.3%; Pred. No. 0.0029;  
Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;  
1334 TCATCACCAGCAGCACACAAAGTTGCTGTTGGAGGGCCAGCGGCACACTCAAGT 1393  
1294 TCCTGACTGGGACCAATCAGACGTACATGGCTGTCAGGGGAGCAGCTACCTTCTGT 1353  
1394 GCAAGCCATTGGGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCGCTGG 1453  
1354 GCAAGGCCCTTCGGAGCGCTGTGCCAGTGTTCAGTGTGACGAGATGGACACAG 1413  
1454 TAGGGAACCTCCTCAAGGACCGCTGTCTATGACAATGGCACCTTGACATCTTCATCACA 1513  
1414 TGCTTCAGACGACGCTTCTCCCTATGCCAATGGGACCTTCGAGACCTCC 1473  
1514 CATCTCAGACAGTGTGCTTCACTGCAATGTGTCGCAATGCTGCCGAGAGCCAGG 1573  
1474 AGGCCAATGACACCGGAGCTACTTCTGCTGGCTGCCAATGACCAAAACAATGTTACCA 1533  
1574 CATGTGGAGGTCTCCATGCTCCAGCTGCCACACCTCAGCAACAGCACAGCCGAC 1631  
1534 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGGCCCCGAGCAC 1591

RESULT 14  
US-08-866-757-1  
; Sequence 1, Application US/08866757  
; Patent No. 5858716  
; GENERAL INFORMATION:  
; APPLICANT: ELISHOURBAGY, NABIL A  
; APPLICANT: LI, XIAOTONG  
; APPLICANT: BERGSMAN, DEBK J  
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RATNER & PRESTIA

ATTORNEY/AGENT INFORMATION:  
NAME: Minnich, Richard J.  
REGISTRATION NUMBER: 24,175  
REFERENCE/DOCKET NUMBER: CWR 2 149-3-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (216) 861-5582  
TELEFAX: (216) 241-1666  
TELEX: (216) 980162  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3774  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acids  
HYPOTHETICAL: irrelevant  
ANTI-SENSE: no  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
INDIVIDUAL ISOLATE: 17-18 week fetus  
IMMEDIATE SOURCE:  
LIBRARY: Stratagene cDNA Library 936206  
CLONE: synthesis of 4 clones  
PUBLICATION INFORMATION:  
AUTHORS: Hlavin, Mary Louise  
AUTHORS: Lemmon, Vance  
TITLE: Molecular structure and functional testing of  
TITLE: human L1CAM: an interspecies comparison.  
JOURNAL: GENOMICS  
VOLUME: 11  
ISSUE:  
PAGES: 416-423  
DATE: 1991  
RELEVANT RESIDUES IN SEQ ID NO: 1 to 3774.  
US-08-427-497E-2

Query Match 1.8%; Score 56.4; DB 2; Length 3774;  
Best Local Similarity 49.3%; Pred. No. 0.0029;  
Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;  
1334 TCATCACCAGCAGCACACAAAGTTGCTGTTGGAGGGCCAGCGGCACACTCAAGT 1393  
1283 TCCTGACTGGGACCAATCAGACGTACATGGCTGCCAGGAGCAGCTACCTTCTGT 1342  
1394 GCAAGCCATTGGGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCGCTGG 1453  
1343 GCAAGGCCCTTCGGAGCGCTGTGCCAGTGTTCAGTGTGACGAGATGGACACAG 1402  
1454 TAGGGAACCTCCTCAAGGACCGCTGTCTATGACAATGGCACCTTGACATCTTCATCACA 1513  
1403 TGCTTCAGACGACGCTTCTCCCTATGCCAATGGGACCTTCGAGACCTCC 1462  
1514 CATCTCAGACAGTGTGCTTCACTGCAATGTGTCGCAATGCTGCCGAGAGCCAGG 1573  
1463 AGGCCAATGACACCGGAGCTACTTCTGCTGGCTGCCAATGACCAAAACAATGTTACCA 1522  
1574 CATGTGGAGGTCTCCATGCTCCAGCTGCCACACTCAGCAACAGCACAGCCGAC 1631  
1523 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGGCCCCGAGCAC 1580

RESULT 13  
US-08-506-296B-13  
; Sequence 13, Application US/08506296B  
; Patent No. 6313265  
; GENERAL INFORMATION:  
; APPLICANT: Phillips, Greg  
; APPLICANT: Cunningham, Bruce A.  
; APPLICANT: Crossin, Kathryn L.  
; TITLE OF INVENTION: NEURITE OUTGROWTH-PROMOTING POLYPEPTIDES  
; TITLE OF INVENTION: CONTAINING FIBRONECTIN TYPE III REPEATS AND METHODS OF USE  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:

STREET: P.O. BOX 980  
CITY: VALLEY FORGE  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICANT: BERGSMAN, DEREK J  
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)  
FILE REFERENCE: GH-70055-1  
CURRENT APPLICATION NUMBER: US/09/153,593A  
EARLIER FILING DATE: 1998-09-15  
EARLIER FILING DATE: 1997-05-30  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 4203

Query Match 1.6%; Score 51.6; DB 2; Length 4203;  
Best Local Similarity 54.1%; Pred. No. 0.033;  
Matches 105; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 684 CAACATGACGGGCTGGTGGACCTGACCCCTGTCCAGGAACACCATCAGCCATCCAGCC 743  
DB 85 CAATCTGCCACCCCTACAGCGCTGACCCCTGTCTCAACAAGATCTCAAGCATCCCTGA 144  
QY 744 CTTTTCCTTTCTGGACCTCGAGAGCTCGGCTCCCTGTCATCTTGACAGCAATCGGCTGCC 803  
DB 145 CTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGTCATCTTCATAACAATAAATTAG 204  
QY 804 AAGCCTTGGGAGGACACCTCCGGGCTGTCTCAACCTGCAGCACCTTATCGTGAACA 863  
DB 205 AAGCCTGAGTCAACACTGTTTGTGATGAGTAACTGGAGACCTTAGACTTTGAATTA 264  
QY 864 CAACAGCTGGGCG 877  
DB 265 TAATAACTTGGGG 278

Search completed: June 8, 2003, 19:06:49  
Job time : 237 secs

STREET: P.O. BOX 980  
CITY: VALLEY FORGE  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICANT: BERGSMAN, DEREK J  
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)  
FILE REFERENCE: GH-70055-1  
CURRENT APPLICATION NUMBER: US/09/153,593A  
EARLIER FILING DATE: 1998-09-15  
EARLIER FILING DATE: 1997-05-30  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 4203

Query Match 1.6%; Score 51.6; DB 2; Length 4203;  
Best Local Similarity 54.1%; Pred. No. 0.033;  
Matches 105; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

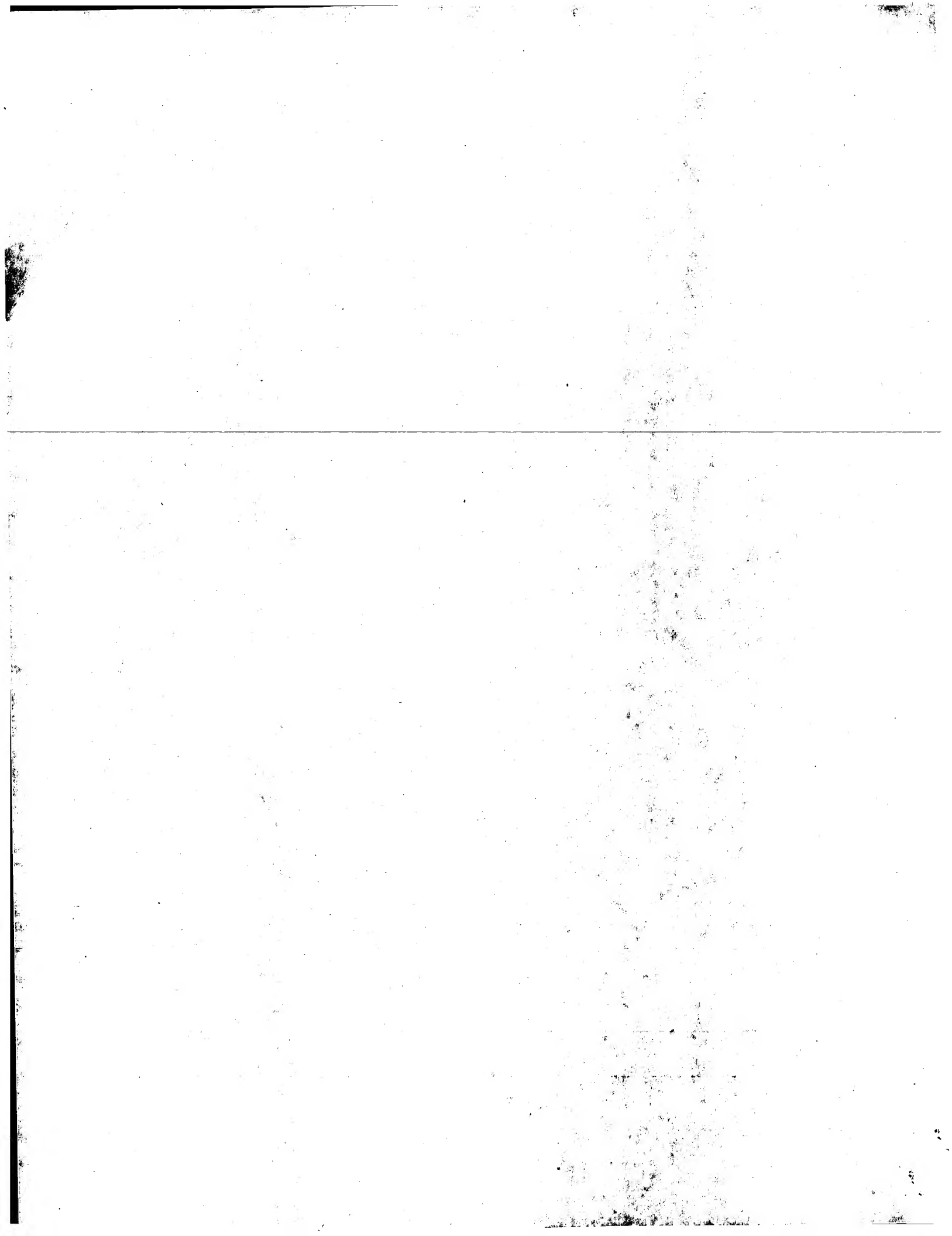
QY 684 CAACATGACGGGCTGGTGGACCTGACCCCTGTCCAGGAACACCATCAGCCATCCAGCC 743  
DB 85 CAATCTGCCACCCCTACAGCGCTGACCCCTGTCTCAACAAGATCTCAAGCATCCCTGA 144  
QY 744 CTTTTCCTTTCTGGACCTCGAGAGCTCGGCTCCCTGTCATCTTGACAGCAATCGGCTGCC 803  
DB 145 CTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGTCATCTTCATAACAATAAATTAG 204  
QY 804 AAGCCTTGGGAGGACACCTCCGGGCTGTCTCAACCTGCAGCACCTTATCGTGAACA 863  
DB 205 AAGCCTGAGTCAACACTGTTTGTGATGAGTAACTGGAGACCTTAGACTTTGAATTA 264  
QY 864 CAACAGCTGGGCG 877  
DB 265 TAATAACTTGGGG 278

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4203 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-866-757-1

Query Match 1.6%; Score 51.6; DB 2; Length 4203;  
Best Local Similarity 54.1%; Pred. No. 0.033;  
Matches 105; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 684 CAACATGACGGGCTGGTGGACCTGACCCCTGTCCAGGAACACCATCAGCCATCCAGCC 743  
DB 85 CAATCTGCCACCCCTACAGCGCTGACCCCTGTCTCAACAAGATCTCAAGCATCCCTGA 144  
QY 744 CTTTTCCTTTCTGGACCTCGAGAGCTCGGCTCCCTGTCATCTTGACAGCAATCGGCTGCC 803  
DB 145 CTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGTCATCTTCATAACAATAAATTAG 204  
QY 804 AAGCCTTGGGAGGACACCTCCGGGCTGTCTCAACCTGCAGCACCTTATCGTGAACA 863  
DB 205 AAGCCTGAGTCAACACTGTTTGTGATGAGTAACTGGAGACCTTAGACTTTGAATTA 264  
QY 864 CAACAGCTGGGCG 877  
DB 265 TAATAACTTGGGG 278

RESULT 15  
US-09-153-593-1  
Sequence 1, Application US/09153593A  
Patent No. 6174994  
GENERAL INFORMATION:  
APPLICANT: ELSHOUBAGY, NABIL A  
APPLICANT: LI, XIAOTONG  
APPLICANT: BERGSMAN, DEREK J  
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)  
FILE REFERENCE: GH-70055-1  
CURRENT APPLICATION NUMBER: US/09/153,593A  
CURRENT FILING DATE: 1998-09-15  
EARLIER FILING DATE: 1997-05-30  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 4203





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 18:21:58 ; Search time 411 Seconds  
(without alignments)  
10705.950 Million cell updates/sec

Title: US-09-831-846-1  
Perfect score: 3144  
Sequence: 1 gcttggtccctctcgtga.....gcgggttgattgattg 3144  
Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues  
Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications, NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	633	20.1	1887	10	US-09-815-626-3
2	633	20.1	2558	10	US-09-815-626-3
3	621.8	19.8	2855	9	US-10-245-103-99
4	621.8	19.8	2855	9	US-10-245-107-99
5	621.8	19.8	2855	9	US-10-245-143-99
6	621.8	19.8	2855	9	US-10-245-171-99
7	621.8	19.8	2855	9	US-10-245-851-99
8	621.8	19.8	2855	9	US-10-245-883-99
9	621.8	19.8	2855	9	US-10-237-535-99
10	621.8	19.8	2855	9	US-10-238-183-99
11	621.8	19.8	2855	9	US-10-238-283-99
12	621.8	19.8	2855	9	US-10-238-370-99
13	621.8	19.8	2855	9	US-10-245-055-99
14	621.8	19.8	2855	9	US-10-245-147-99
15	621.8	19.8	2855	9	US-10-245-730-99
16	621.8	19.8	2855	9	US-10-245-739-99
17	621.8	19.8	2855	9	US-10-246-210-99
18	621.8	19.8	2855	9	US-10-239-196-99
19	621.8	19.8	2855	9	US-10-243-024-99

20	621.8	19.8	2855	9	US-10-243-409-99	Sequence 99, Appl
21	621.8	19.8	2855	9	US-10-245-033-99	Sequence 99, Appl
22	621.8	19.8	2855	9	US-10-245-621-99	Sequence 99, Appl
23	621.8	19.8	2855	9	US-10-245-880-99	Sequence 99, Appl
24	621.8	19.8	2855	9	US-10-243-095-99	Sequence 99, Appl
25	621.8	19.8	2855	9	US-10-245-185-99	Sequence 99, Appl
26	621.8	19.8	2855	9	US-10-245-427-99	Sequence 99, Appl
27	621.8	19.8	2855	9	US-10-245-473-99	Sequence 99, Appl
28	621.8	19.8	2855	9	US-10-245-770-99	Sequence 99, Appl
29	621.8	19.8	2855	9	US-10-245-877-99	Sequence 99, Appl
30	621.8	19.8	2855	9	US-10-246-976-99	Sequence 99, Appl
31	621.8	19.8	2855	9	US-10-243-320-99	Sequence 99, Appl
32	621.8	19.8	2855	9	US-10-243-743-99	Sequence 99, Appl
33	621.8	19.8	2855	9	US-10-242-845-99	Sequence 99, Appl
34	608	19.3	1653	10	US-09-897-214-9	Sequence 9, Appl
35	554.6	17.6	1653	10	US-09-897-214-9	Sequence 9, Appl
36	477.2	15.2	8269	9	US-10-125-540-551	Sequence 551, App
37	477.2	15.2	8269	9	US-10-073-961-452	Sequence 452, App
38	477.2	15.2	8269	9	US-10-103-313-537	Sequence 537, App
39	477.2	15.2	8269	10	US-09-764-887-452	Sequence 452, App
40	477.2	15.2	8269	10	US-09-764-870-551	Sequence 551, App
41	477.2	15.2	8269	10	US-09-764-853-879	Sequence 879, App
42	461.4	14.7	1872	10	US-09-768-826-21	Sequence 21, Appl
43	273.6	8.7	493	9	US-09-918-995-26236	Sequence 26236, A
44	204.2	6.5	432	10	US-09-897-214-4	Sequence 4, Appl
45	204.2	6.5	432	10	US-09-893-737-189	Sequence 189, App

ALIGNMENTS

RESULT 1

US-09-815-626-3  
; Sequence 3, Application US/09815626  
; Patent No. US20020076752A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria Alexandra  
; TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT  
; FILE REFERENCE: 10448-031001  
; CURRENT APPLICATION NUMBER: US/09/815,626  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/191,863  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1887  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-815-626-3

Query Match	20.1%	Score	633;	DB	10;	Length	1887;		
Best Local Similarity	61.3%	Pred. No.	9e-166;						
Matches	1078;	Conservative	0;	Mismatches	665;	Indels	16;	Gaps	3;
QY	507	GTTCGCGTGGTGCAGCGCTGCCCAAGTACTGTCTGTCGCCAGAAATCTGTCTGAGTCACT	566						
Db	63	GTACGCCACACCCAGCCCATGTCCCGCGGCTGCCGTCGCCAGACACAGTCGCTGCCCT	122						
QY	567	GGGGACCTGTGTCCTTCAAGGGGCTGCTCTTTGTACCCCTGATATTGACCGGGGAC	626						
Db	123	AAGCGTGTGTGCCAGGGGAGGCTCTCTGTGTGTCACCTCGCTGGACCGCGGGC	182						
QY	627	AGTGAGCTGCGCTGGGGGCACTTCATCATCCACATCAGCGCGCCAGGACTTTGCCAA	686						
Db	183	AGCCGAGCTCGGCTGGCAGACAACTTCATCGCTCCGTGCGCGCCGACCTGGCCAA	242						
QY	687	CATGACGGGCTGGTGGACCTTGACCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT	746						
Db	243	CATGACAGGCTCTGTCATCTGAGCTGTGCGGGAACACCATCCGCCAGTGGCTGCCGG	302						
QY	747	TTCTTTTGGACCTCGAGAGCCTCCGCTCCCTGTCATCTTGACAGAATCGGCTGCCAAG	806						

Db		1375	ATCCAGTACAA	CAGCTCGGCTGATGACATCCTTGTTACAGAGTAGATCCC	GGCGGAGACC	1433
Qy		1885	AAGGCCTTCGTGGTCAACA	AACTGGTGTACAGGACTGGCTACGA	CTTGTTGTGCTGGCC	1944
Db		1435	CGCTCGTTCCTGCTGACG	GAACTCTGGGTACAGGCCGACCTACGATCTGTGCGTGTCTCGCC	1494	
Qy		1945	ATGTGGGATGACACAGCAC	ACATCTCAGGGCCACACATCGTGGGCTGGCCCCAGTTCTC	2004	
Db		1495	GTGTATAGGGCAGTGCCA	CGGGGCTCAGGGCCAGCGGCTGTGGGTGTGGCCCGCTTTC	1554	
Qy		2005	TTCACCAAGCTGACTACT	CCCGCAGTGCACATCCATGCACAGCCAGATTCTGGGCGGCA	2064	
Db		1555	TCACACGAACCTGGCTG	CGGCATCGGGGGCCGCACGCTCCCTTCTCTGGGCGSCAG	1614	
Qy		2065	ATGATCCTGTTCATCGGG	GGGCATCATCTGGGCGGCATCTGTTGGCCACGCTGCTGGTCTTCATCTCATCTCTC	2124	
Db		1615	ATGATCATCGCTGGCG	GGTTCATCTGAGCTTCGTTACTTTCGTTGCTG	1674	
Qy		2125	ATGGTGGCTTACAAGGT	CTGCAACCAAGGAGGCCCCCAG-----CAAGATGGCAGCGGCC	2178	
Db		1675	CTAATGGCTTACAAGST	GCACGGCGGCGAGCCCCCGGCAAGCTTC	1734	
Qy		2179	GTGAGCAATGTACTCC	GAGAACAGGGCGCCAGCCACCGGCTCTCAAGCAGCGCACCA	2238	
Db		1735	GTTAGCAGCTTTCCT	CCAGACCAAGCGGCGCTTGGGCCCCACGCCCCG	1794	
Qy		2239	GCCGGGCGCCGCGCAGG	2257		
Db		1795	CCGCCCCGGAGCCCGCG	1813		
 RESULT 2 US-09-815-626-1 ; Sequence 1, Application US/09815626 ; Patent No. US20020076752A1 ; GENERAL INFORMATION: ; APPLICANT: Glucksmann, Maria Alexandra ; TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT ; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF ; FILE REFERENCE: 10448-031001 ; CURRENT APPLICATION NUMBER: US/09/815,626 ; CURRENT FILING DATE: 2001-03-23 ; PRIOR APPLICATION NUMBER: US 60/191,863 ; PRIOR FILING DATE: 2000-03-24 ; NUMBER OF SEQ ID NOS: 17 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 1 ; LENGTH: 2558 ; TYPE: DNA ; ORGANISM: Homo sapiens ; FEATURE: ; NAME/KEY: CDS ; LOCATION: (435)...(2318) US-09-815-626-1						
 Query Match                      20.1%;    Score 633;    DB 10;    Length 2558; Best Local Similarity         61.3%;    Pred. No. 9.9e-166; Matches 1078;    Conservative         0;    Mismatches 665;    Indels    16;    Gaps         3						
Qy		507	GTTTGGCTGGTGCAGCG	CGCTGCCCCAAGTACTGTCTGCCAGAATCTGTCTGAGTCACT	566	
Db		497	GTCAAGCACACCAGCC	CCATGTCCCCCGCGCTGCCGACACAGTCTGCTGCCCT	556	
Qy		567	GGGACCCCTGTGCCCT	TCCNAGGGGTGCTCTTTGTACCCCTGATTTGACCGCGCGGAC	626	
Db		557	AAGCGCTGTGTGCC	CCAGGGCAGGCGCTCTGTTCGTGCCACCTCGCTGGACCCCGGC	616	
Qy		627	AGTGGAGCTCGGCTG	GGCGGCAACTTCATCATCCACATCAGCGCGCAGGACTTGC	686	
Db		617	AGCGAGCTCGGCTG	SGCACAACTTCATCGCTTCGTCGCGCGCGGCACTGGCC	676	
Qy		687	CATACGCGGCTGTG	ACCTGACCTGCTCCAGGAACACCATCAGCCACATCCAGCC	746	

Db	303	CGCCTTGC	CGACCTTGC	CGGCGCCTCGGTGCCTCGACCTGATGC	CAACCGGTGACCTC	362
Qy	807	CTTTGGGG	AGGACACCTCGGGGCTGGTCAACCTGCAG	CACCTTATCGTGAACAA	CAAA	866
Db	363	ACTGGG	CGAGGGCAGCTCGGGGCTGGTCAACTTGGCC	ACCTCATCTCAGCA	CAAA	422
Qy	867	CGAGCT	GGGGGCGATCGCAGATGAGGCTTTTGAGGACTTCTGTG	CAATTTGGAGGATCT		926
Db	423	CGAGCT	GGCAGCCTCGCGCGCGGCGCTTGATGATGTCG	GAGACACTGGAGAC		482
Qy	927	GGACCT	CTCTCAACAACTCATGCGCTGCGGTGGGACTCCGTG	CGACGCATGTG	CAAA	986
Db	483	CGACCT	CTCTCAACAACTCAGCAGCTGCGTGGGAGGCCCTTGG	CGCCGCTTGGGCA		542
Qy	987	CTCCAC	CAGCTGAGCTGACCAACCTGTGTGATCACATCG	CCGAGGACACTTTGC		1046
Db	543	CGTCAA	CACGTTGGGCTTCGACCAAACTGCTGGCTTCTGTG	CCGCGCGGCTTTTTC		602
Qy	1047	AGACCT	GCAGAAATGCGCCGCTTGATCTACCTCCAATCGGCTG	CAGAGCTGC		1106
Db	603	CCGCTG	CACAAGCTGGCCCGGTGGACATGACTCAACCCGCTG	ACCAATCCGACC		662
Qy	1107	TGATCC	CACTTTTCCCGCTCCAGGCTTCGGCTTTGACAGCA	CAACCTTTTGCCCA		1166
Db	663	CGACCA	CTCTTCTCCCGCTGCGCTGCTCGACGCGCCGGGCT	CGCCGCTCTGC		722
Qy	1167	CTTTGCT	TTTAACTTTGGGGTAAACCACTTCACTGCAATTTG	GAGCTTCTTGCTGG		1226
Db	723	CTTGTG	CTGGCTTTTGGCGGAACCCCTGCACTGCACTG	CAGACTGGTGTGCTGCG		782
Qy	1227	GAGGCT	CGAGCGGACGATGACTGGAACCTGTGGCTCCCA	AGGGGCTCAAGGGTGC		1286
Db	783	TCGCTG	CGCGGGAGGACCTTCGAGGCTCGCGCTCCCA	CTGCTCGGGCGCG		842
Qy	1287	CTACTT	CTTGCAATGTGCTGAGGAGTTTGTGTGCGAG	CGCGCTCTCATCCCGCA		1346
Db	843	CTACTT	TTTGGCGGTGGCGAGGAGTTTGTGTGCGAG	CGCGCGGTGTGACTCACCG		902
Qy	1347	CACACA	AAAGTTGTGTTTGGAGGCGCAGCGGCGCACACT	CAAGTGC	AAAGCCATTGG	1406
Db	903	CTCAC	CACCTCTGCGTGGCGCGCAGCTGCGCGGCTG	CGCTCGCGG	CAGTGGG	962
Qy	1407	GGACCC	AGCCGCTTATCCACTGGGTAGCCCCGATGACG	CGCTGTGAGGAACTCCTC		1466
Db	963	GGACCC	CAGAGCCCCGTGTGGTTGTCACCCCAGGCG	CGGTGTAGGCAACTCAAG		1022
Qy	1467	AAGGA	CCGCTGTATGACAATGGCACCTTGGACATCTT	CATCACCACATCTCAGGACAG		1526
Db	1023	CCGTG	CCCGCGCTCCCAATGGGACGCTGAGGCTGCTG	GTCAACGCGGGTATGG		1082
Qy	1527	TGCTG	CTTCACTGATTTGCTGCCAATGCTCGCGAGAG	CGCCACGCGCATGTGTGAGGT		1586
Db	1083	TGGCAT	CTTCACTGATTTGGGCCAATGACGTGGCGG	CCAAATGACGTGGAGGCCACAGCTGTGTGTG	AGACT	1142
Qy	1587	CTCAT	GTGCCAGCTGCCACCTCAGCAACAGCACAG	CGCGCA--CTGCACCC	CCCCCAAG	1644
Db	1143	GACTGT	GGGTCCCCCAACCACTCTCAGTAGCCAA	CAGCACAGCTGTGACCCCGCG		1202
Qy	1645	TCCG	CGCTCTCAGACATCACTGGCTCCAGCAAGA	CCAGCGGGAGGTGGAGG	CAGTGGG	1704
Db	1203	GGAC	GGGATCTGTATGCTCTCACCCACCTCCGCT	CTGCTTCTGCAAGGTGGC		1262
Qy	1705	GGCG	GAGGCTCCCAAAAGCCCCCGGAACGGGCTGT	GCTTGTCTGTGAAGTGAC	CACACC	1764
Db	1263	CGACA	CTGGGCCCTTA-----CCGACCGTGGCGT	CCAGGTGACTTGAGCA	CGGGGCC	1314
Qy	1765	ACCT	CGGCGCTGTGCAAGTGGTCTGTGACCAAGT	CAGCACCCCGGGTGAAGATGTAC	ACAG	1824
Db	1315	ACAG	TGCTTGTTCAGTGGCGGATCAGCGGCTAT	CCCGGCAATCCGATGTAC	ACAG	1374
Qy	1825	CTGC	AGTCAACTGCTCTGACGATAGGTA	CTGATTTACAGGATGATCCAG	CTCCAC	1884

1825	Qy	CTCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCAGCCTCCAAC	1884
1809	Db	ATCCAGTACAAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC	1868
1885	Qy	AAGGCGCTTCGTGCTCAACAACTGGTGTCAGGGACTGGCTACAGACTTGTGTGCTGTGGCC	1944
1869	Db	CGTCTGTTCTGTCTGACGGACCTGGCGTCNAGGCCGGACCTACCATCTGTGCGTGTCTGCC	1928
1945	Qy	ATGTGGGATGACACAGCAGCAGCACTCACGGGCCACCAACATCTGTGGGCTCGGCCCGCAGTTTC	2004
1929	Db	GTGTATGAGGGCAGTGCACGGGGCTCAGCGGCACCGCGCTCTGGGCTCGGCCCGCGTTTC	1988
2005	Qy	TTCAACAAAGCTCTACTACCCGCGAGTGCAGTCTCATGCACAGCCAGATTTCTGGCGGCGCAC	2064
1989	Db	TCACCGAACTCTGGCTGCGGCCATCGGGGGCGCGCAGCTCCCTTCCTGGGGGGCAGC	2048
2065	Qy	ATGATCCTGTGTCATCGGGGGCATCATCTGTGGCCACGCTGCTGGTCTTCATCGTCACTCTC	2124
2049	Db	ATGATCATCGCTGGCGGCGCTCATCTGATGCCCTCGGTACTGGTCTTCATCTTCGTGCTG	2108
2125	Qy	ATGGTCGGCTACAAGTCTCTCAACACAGGAGGCCCGCCAG-----CAAGATGGCAGCGGCC	2178
2109	Db	CTAATGGCTTACAAGGTGACGGCGGCCAGCCCCCGGCAAGGCCAAGATTCCCGCGGCC	2168
2179	Qy	GTGAGCAATGTGTACTCGCAGACCAACGGCGGCCCGCAGCACCGCTCTCAAGCAGCGCACCA	2238
2169	Db	GTTAGCAGCGTTTGTCTCCACAGCAACAGCGCGCCCTGGGCCCCACGCCACCGCCGCCCG	2228
2239	Qy	GCCGGGGCCCCCGCGCAGG	2257
2229	Db	CCGCGCCCGGAGCCCGCGG	2247

### RESULT 3

US-10-245-103-99

Sequence 99, Application US/10245103  
Publication No. US20030068778A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin  
APPLICANT: Eaton, Dan  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Goddard, Audrey  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin  
APPLICANT: Smith, Victoria  
APPLICANT: Stephan, Jean-Philippe  
APPLICANT: Watanabe, Colin  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
APPLICANT: Fong, Sherman  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
ACIDS ENCODING THE SAME  
FILE REFERENCE: P3630R1C112  
CURRENT APPLICATION NUMBER: US/10/245,103  
CURRENT FILING DATE: 2002-09-17  
PRIOR APPLICATION NUMBER: 10/157942  
PRIOR FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/059114  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/063046  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/065027  
PRIOR FILING DATE: 1997-11-10  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/086478  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/090557

```

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
; LENGTH: 2855
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-103-99

Query Match      19.8%; Score 621.8; DB 9; Length 2855;
Best Local Similarity 61.3%; Pred. No. 1.3e-162;
Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

Qy 507 GTTTGCGGTGGTGAAGCCCTGCCCCAAGTACTGTCTGCGCAGATCTGTCTGAGTCACT 566
Db 798 GTCAGGCACACCCAGGCCATGTCCCGCGCTGCCGCTGCCAGACACAGTCTGCTGCCCT 857

Qy 567 GGGGACCCCTGTGCCCTCCAAAGGGGTGCTCTTTGTACCCCCCTGATATTGACCGCGGAC 626
Db 858 AAGCGTCTGTGCCAGGGGAGCGCTCTCTGTTCGTGCACCTCGCTGGACCGCGGGC 917

Qy 627 AGTCGAGCTGCGCTTGGCGGGCAACTTCATCATCTCACATCAGCGCGCAGGACTTTGCCAA 686
Db 918 AGCGAGCTGCGGTGCGACACAATTCATCGCCTCCGTGCGCGCGGCGACTTGGCCAA 977

Qy 687 CATGACGGGCTGTGACCTGACCTGTGCCAGGAACACATCAGCCACATCCAGCCCTT 746
Db 978 CATGACAGGCTCTGTGATCTGAGCTGTGCGGAACACCATCGCCACGTGGTGGCGG 1037

Qy 747 TTCCTTTCTGACCTTCGAGAGCCTCCGCTCCCTGCAATCTTGACGAATTCGCTGCCAAG 806
Db 1038 CGCCTTCGCCACCTCGGGGCCCTGCGTCCCTGCACCTGGATGGCAACCGCTGACCTC 1097

Qy 807 CTTGGGAGAGACACCTTCGGGGCTGTCAACTCGACCTCGACACCTTATCGTGAAACAACAA 866
Db 1098 ACTGGCGAGGGCCAGCTGCGCGGCCCTGGTCAACTTGGCGCACCTCATCTCAGCAACA 1157

Qy 867 CCAGCTGGCGGCGATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTCGAGGATCT 926
Db 1158 CCAGCTGGCAGGGCTGGCGGGCGCGGCCCTGGATGATTGTGCCGAGACATCGGAGACCT 1217

Qy 927 GGACCTCTCTTAACAACAACCTTCATGCGCTGCGCTGCGGACTTCGTGGCAGCATGTCAA 986
Db 1218 CGACCTCTCTTAACAACAACCTCGAGAGCTGCCCTGGGAGGCCCTGGGCGCGCTGGGCAA 1277

Qy 987 CTTCCACAGCTGAGCTGGACCAACAACCTCTGGAATCACATCGCGAGGGCACCTTTGCG 1046
Db 1278 CGTCAACAGTTTGGGCTCGACCAACAACCTCTGGCTCTGTGCTGCTGCTGCTGCTGCTG 1334

Qy 1047 AGACCTGCAGAACTGGCGCGCTGGATCTCACCTCCAATCGGTGCGAAGCTGCCGCC 1106
Db 1335 CGGCTGCACAGCTGGCCCGGCTGAGATGACCTCCAACCGCTGACCAACAATCCACCC 1394

Qy 1107 TGAATCCATCTTTTGGCGCTCCAGGCTTCGCTTTGACGACACACCTTTTCGCCCAACC 1166
Db 1395 CGACCCACTCTTCTCCCGCTGCGCTGCGCAGGCCCGGGGCTGCGCGCTCTGCG 1454

Qy 1167 CTTGTCTTTAGTTTGGGGGTAAACCACTTCACCTGCAATTTGTGAGCTTCTCTGGGTGG 1226
Db 1455 CTTGTGCTGGCTTTTGGCGGNAACCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy 1227 GAGGCTCGAGCGGACGATGACCTGGAACCTGTGGCTCCCGAGGGGCGCTCAAGGGTGC 1286
Db 1515 TCGCTTGGCGCGGAGGACGACCTCGAGGCTGCGCGTCCCGACCTGCTCTGGGCGGCGG 1574

Qy 1287 CTACTTCTGCATCTGCGTCAGAGAGTTTGTGTGCGAGCGCGCTCTCATCACCCAGCA 1346
Db 1575 CTACTTCTGGCGGTGGCGAGAGAGTTTGTCTGCGAGCGCGCGTGGTGACTCAACCG 1634

Qy 1347 CACACACAAGTTGCTGTTCTGGAGGGCCAGCGGCCACACTCAAGTGCAAAGCCATGG 1406

```

1218	CGACCTCTCTCTACAAACCTCGAGAGAGTGCCCTGGAGAGCCCTGGCGCGCTGGCGAA	1277
987	CTCTCCACCAAGCTGAGCCTGGACACCAACCTGCTGGATCACATCGCCGAGGACACTTTGC	1046
1278	CGTCAACAGTTGGGCTCGACCAACACTGCTGGCTTCTGTGCC--GGCGCTTTTTC	1334
1047	AGACTGCGAGAACTGGCGCGCTGGATCTCACTTCCAATCGGCTCGAGAACTGGCCCC	1106
1335	CGCGCTGCACAAGCTGGCGCGCTGGAGATGACCTTCCAAACGCGCTGACCAACAATCCCA	1394
11107	TGATCCCATCTTTGGCCGCTCCAGGCTTCGGCTTTGACGCCACACCTTTGGCCCCACC	1166
13195	CGACCCACATCTTCTCCCGCTCGCCCTGCTGCGAGGCCCGGGGCTCGCCGCGCTCTGC	1454
11167	CTTGTCTTTTGTGTTTTGGGGGTAAACCACTTCACTGCAATTTGTGAGCTTCTCTGGCTGC	1226
14155	CGTGTGCTGGCTTTGGCGGGAACCCCTGCATGCAACTCGAGCTGGTGTGCTGCTGCG	1514
1227	GAGGCTGAGCGGACGATGACCTCGAAACCTGTGGCTTCCCGAGGGGCTCTCAAGGCTCG	1286
15151	TCGCTGGCGCGGAGGACGACTCGAGGCTTCGCGGTCCCGTCCACACTGCTCTGGCGGCG	1574
1287	CTACTTCTGCGATGTCGCTGAGGAGGTTTTGTGTGCGAGCGCGCTCTCATCACCCAGCA	1346
15175	CTACTTCTGGCGGTGGCGAGGAGGTTGTGTGCGAGCGCGCGGTGTGTTGACTCAACG	1634
1347	CACACAAAGTTGCTGGTTCTGAGGGGCGAGCGGCCACACTCAAGTGCMAAGCCATTGG	1406
1635	CTACACACTCTGGCTGCGCGAGCTGGCGGCTGCCCTGCGCTCGGGGAGTGGG	1694
1407	GGACCCAGCGCCCTTATCCACTGGGTAGCCCCCGATGACCGCTGTGTGAGGAACTCCTC	1466
1695	GGACCCAGAGCCCGTGCCTGTGGGTGTCAACCCAGGGCGGCTGTAGGCAACTCAAG	1754
1467	AGGACCGCTGCTATGACAATGGACACCCTGGACATCTTCATCACCACATCTCAGGACAG	1526
171755	CGGTGCGCGCGCTTCCCAATGGGACGCTGGAGTGTGTGTACCGAGCCGGGTGATGG	1814
1527	TGTTGCTTCACCTGATTTGCTGCAATGCTGCCGGAGAGCCAGCGCATGTGTGGAGGT	1586
1815	TGGCATCTTCACTGATTTGCGGCCAATGCACTGGCAGAGGCCACAGCTGCTGTGGAGCT	1874
1587	CTCATGCTCCAGCTGCCACACTCAGCAACAGCACACGACGCA--CTTGACACCCCAAG	1644
1875	GACTGTGGTCCCCACCACTCTCAGCTAGGCAACAGCACAGCTGTGACCCCGCGG	1934
1645	TCGGGCTCTCAGACATCACTGGCTCCAGCAGACACGCGGGGAGGTGAGGACAGTGGG	1704
1935	GGACGGGATCTGTATGTCTCAACCCACCTTCGCCCTGCCTCTGCTTCTGCAAGGTGGC	1994
1705	GGCGGAGAGCTCCCAAAAGCCCCCGAAAGGGCTGTGTGTCTGTGAAGTGACCAACC	1764
1995	CGACATGGGCGCCCTA-----CCGACCGTGGCTCAGGTGACTGACACGGGGCC	2046
1765	ACCTCGGCGCTGTCAGTGGTGTGTGCAAGATCAGACACCCCGGGTGAAGATGTACCA	1824
2047	ACAGCTGCTCTTGTCCAGTGGCGGATCAGCGGCTATCCCGGGCATCCCGATGTACCA	2106
1825	CTGCAGTAACTTGCTCTGACAGATGAGTATCTGATTTACAGATGATCCAGCCTCCAAC	1884
2107	ATCCAGTACAAAGCTCGGTGATGACATCTCTGCTTACAGGATGATCCCGGGGAGAGC	2166
1885	AAGGCGCTTGTGTGTCAAACACTGTTGTGAGGACTGGCTACGACTGTGTGTGCTGCGC	1944
2167	CGCTGTTCTCTGTGACGAGCTTGGGCTCAGGCCGACCTACGATCTGTGGTGTCTCGCC	2226
1945	ATGTGGGATGACACAGCCACGACTCTACGGGCCACCAACATCTGTGGGCTGCGCCAGTTC	2094
2227	GTGTATGAGACAGCCACCGGGCTCAGGGCCACGCGGCTGTGGGCTGTGGCGGCTTC	2286
2005	TTCAACAGGCTGACTACCGCAGTGCCAGTCCATGCAACAGCCAGATTCTGGGCGGCAAC	2064

Tue Jun 10 08:57:37 2003

Db 2287 TCCACGAACTGCGCTGGGCGCATGCGGGGCGCGACGCTCCCTTCTGGGCGGACG 2346  
 QY 2065 ATGATCCTGGTCAATCGGGGGCATATCGTGGCCAGCGCTGCTGCTTTCATGTCATCTC 2124  
 Db 2347 ATGATCATCGCTGGGCGGCTCATCGTAGCTCGTACTGGTCTTTCATCTTCTGCTG 2406  
 QY 2125 ATGGTGGCTACAGGTCTGCAACCAACGAGGCGCCCGCAG-----CAAGATGGCAGCGGCC 2178  
 Db 2407 CTAATCGCTACAGGTGCAACGCGGCGCCAGCCCGCGGCAAGGCGCAAGATTCCCGGCGCT 2466  
 QY 2179 GTGAGCAATGTGTAATCGGAGACCAACGCGGCGCCAGCCCGCTTCCAGAGCGGCACCA 2238  
 Db 2467 GTTAGCAGGCTTGTCTCCAGACCAACGCGGCGCTTGGGCGCCAGCCCGCGCGCG 2526  
 QY 2239 GCCGGGCGCCCGCGGAGG 2257  
 Db 2527 CCGGCGCGGAGCGCGG 2545

RESULT 5  
 US-10-245-143-99  
 ; Sequence 99, Application US/10245143  
 ; Publication No. US20030068780A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stephan, Jean-Philippe  
 ; APPLICANT: Watanabe, Colin  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; APPLICANT: Fong, Sherman  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3630R1C90  
 ; CURRENT APPLICATION NUMBER: US/10/245,143  
 ; CURRENT FILING DATE: 2002-09-16  
 ; PRIOR APPLICATION NUMBER: 10/197942  
 ; PRIOR FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/059114  
 ; PRIOR FILING DATE: 1997-09-17  
 ; PRIOR APPLICATION NUMBER: 60/063046  
 ; PRIOR FILING DATE: 1997-10-24  
 ; PRIOR APPLICATION NUMBER: 60/065027  
 ; PRIOR FILING DATE: 1997-11-10  
 ; PRIOR APPLICATION NUMBER: 60/079689  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/086478  
 ; PRIOR FILING DATE: 1998-05-22  
 ; PRIOR APPLICATION NUMBER: 60/087607  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090689  
 ; PRIOR FILING DATE: 1998-06-25  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 116  
 ; SEQ ID NO 99  
 ; LENGTH: 2855  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 ; US-10-245-143-99

Query Match 19.8%; Score 621.8; DB 9; Length 2855;  
 Best Local Similarity 61.3%; Pred. No. 1.3e-162;  
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

QY 507 GTTTCGCTGCTGACGCGCTGCCCAAGTACTGTGTCTGCGAGAACTCTGTCTGAGTCACT 566  
 Db 798 GTACGCCACACCCAGCCCATGTCCCGCGCTGCGCTGCCAGACACAGTCCGCTGCCCT 857  
 QY 567 GGGGACCCCTGTGCCCTCCAAAGGGCTGCTCTTTGTACCCCTGATATTTGACCGCGGAC 626  
 Db 858 AAGCGTGTGTGCCAGGGGAGGCTCTCTGTTCGTGCACTCTGCTGAGACCGCGGCG 917  
 QY 627 AGTGGAGCTGCGCTGGGCGGCACTTCATCATCCACATCAGCCGCCAGGACTTTTGC 686  
 Db 918 AGCCGAGCTGCGGCTGGCAGCAACTTCATCGCTCCGTCGCGCGCGGCGGCTGCCAA 977  
 QY 687 CATGACGGGCTGTGTGAGCTGACCTGTCAGGAAACCATCAGCCACATCCGCGCTG 746  
 Db 978 CATGACAGGCTGTGCTGATCTGAGCTGTGCGGGAACCATCCTGGATGGCAACCGG 1037  
 QY 747 TCCCTTTCTGAGCTCGAGGCTCGCTCCCTGATCTTGACAGCAATCGGCTGCAAG 806  
 Db 1038 CGCTTTGCGCGACCTCGGGGCTGCTGCTGCTGACCTGGATGGCAACCGGCTGAC 1097  
 QY 807 CCTTGGGAGGACACCTCTCCGGGCTGTGTAACCTGACAGCACCTTATCGTGAACAA 866  
 Db 1098 ACTGGCGAGGCGAGCTGCGCGCTGTGTAACCTGCGCCACCTCATCTCAGCAACA 1157  
 QY 867 CCAGCTGGCGGCTGCGAGATGAGCTTTTGGAGACTTCTGCTGACATTTGAGGATCT 926  
 Db 1158 CCAGCTGGCAGGCTGGCGCGGCTGCTGATGATTTGCGGAGACATCGGAGGACCT 1217  
 QY 927 GGACCTCTCTACAAACCTCTCATGCGCTGCGGCTGCGGACTTCTGCGGAGCAT 986  
 Db 1218 CGACCTCTCTTACAAACCTCTGAGGAGCTGCGCTGGGAGGCTTGGGCGCTGGGCA 1277  
 QY 987 CCTCCACGCTGAGCTGGAACAACCTGCTGGATCATCATCGCGAGGCGACCTTTGC 1046  
 Db 1278 CGTCAACACGCTGGGCTTCGACCAACCTGCTGGCTTCTGTGCCC--GGCGTTTTC 1334  
 QY 1047 AGACTGAGAACTGGCGCGCTGTGATCTCACCTCCAATCGGCTGAGAACTGCCCC 1106  
 Db 1335 CGCGCTGCAAGCTGGCGCGCTGGACATGACCTTCAACCGCTGACCAATCCAC 1394  
 QY 1107 TGATCCCATCTTTGCGCGCTCCAGGCTTCCGCTTTGACAGCACAACCTTTGCC 1166  
 Db 1395 CGACCCACTCTTCTCCGCTGCGCTGCTGCGGAGGCGCGGCGCTGCGCGCTGCG 1454  
 QY 1167 CTTGTCTTTAGTTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTCTGGCTGC 1226  
 Db 1455 CTTGTGCTGGCTTTGGCGGAAACCTCTGCACTGCACTGCGAGCTGCTGCTGCTG 1514  
 QY 1227 GAGGCTCGAGCGGAGCTGACCTGGAAACCTGTGGCTTCCCGAGGGGCTCNAAGG 1286  
 Db 1515 TCGCTGCGCGGAGGAGCTGAGGCTGCGCGCTGCGCTGCGCTGCTGCGGCGCG 1574  
 QY 1287 CTACTTCTGGCTGCTGAGGAGGAGTTTGTGCGAGCGCTCTCTCATCACCAGCA 1346  
 Db 1575 CTACTTCTGGCGGCTTGGCGGAGGAGTTTGTGCGAGCGCGCTGCTGCTGCTG 1634  
 QY 1347 CACACAAAGTTGCTGTTCTGGAGGCGGAGGCGGCACTCAAGTGAAGGAGGCTGG 1406  
 Db 1635 CTACCACTCTGCTGCTGCTGCGCGGCTGCGCGCTGCGCTGCGCTGCGGCGAG 1694  
 QY 1407 GGACCCCGAGCCCTTATCCACTGGGTAGCCCCGATGACCGCTGCTGAGGAGGCT 1466  
 Db 1695 GGACCCCGAGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1754  
 QY 1467 AAGGACCGCTGCTATGACAAATGGCACCTTGGACATCTTATCAACACATCTCAG 1526  
 Db 1755 CCGTGGCGGCGCTTCCCAATGGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCT 1814  
 QY 1527 TGGTGGCTTCACTGAGATCTGCGCAATGCTGCGGAGGCGGCGGCGGCTGCTGCTG 1586  
 Db 1815 TGGCATCTTCACTGCAATGCGGCAATGAGCTGCGGAGGCGGCGGCGGCTGCTGCTG 1874  
 QY 1587 CTCATCTGCTGCGGCTGCGGCACTCAGCAACAGCAGCAGCGGCA--CTGACCCCG 1644

Db 1875 GACTGTGGTCCCCACCACTCTCTCAGCTAGCAACAGCAGCTGTGACCCCGCG 1934  
Qy 1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACAGCGCGGAGGTGGAGGAGTGG 1704  
Db 1935 GGACGGGATCTGTATGCTCTCACCCACCTCCGCTGCTCTCTTCTGCAAGGTGGC 1994  
Qy 1705 GCGGAGAGCTCCCAAAAGCCCCCGGAACGGGTGTGTTGTCTGAAGTGACCAACC 1764  
Db 1995 CGACACTGGGCCCTTA-----CCGACCGTGGCTCCAGGTGACTGACGACGGGCC 2046  
Qy 1765 ACCTCGGCCCTGGTCAAGTGGTCTGTACAGCAAGTACAGACCCCGGGTGAAGTGTACCA 1824  
Db 2047 ACAGTGTCTTGTCTCAGTGGCGGATCATCGCGGCTATCCCGGGATCCGCATGTACCA 2106  
Qy 1825 CTGCACTAAGTCTGACATGAGTACTGATTTACAGATGATCCAGCCTCCAAAC 1884  
Db 2107 ATCCAGTACAGCTGGCTGATGACATCTCTGCTACAGATGATCCCGGGAGAGC 2166  
Qy 1885 AAGGCTTGTGTTCAACAACTGTGTACAGGACTGGCTACGACTTGTGTGTGGCC 1944  
Db 2167 CGCTCGTCTGACGACCTGGCTCAGCGGACCTACGATCTGTGCTGTGCTGCGC 2226  
Qy 1945 ATGTGGGATGACAGCCAGCACATCAGCGGCAACCAATCGTGGGCTGCCCGAGTTC 2004  
Db 2227 GTGTATGAGACAGCGCCAGCGGGCTCACGGCACCGCGGCTGTGGGCTGCGCCGCTTC 2286  
Qy 2005 TTCAACAGGCTGACTACCCGAGTGCAGTGCAGTCCATGSCAGCAGATCTCGGCGGCACC 2064  
Db 2287 TCACCGAACTCGCTGCGGCGCATCGGGGGCGCGCAGCTCCCTTCTGGGGCGCAG 2346  
Qy 2065 ATGATCTGTGTCATCGGGGCGCATCTGTGGGCGAGCTGCTGTCTTTCATCTGTCCTC 2124  
Db 2347 ATGATCATCGCTGCGGGCGCTCATCTGAGCTCGTACTGTCTTTCATCTGCTGTG 2406  
Qy 2125 ATGTTGCGCTACAAGTCTGCAACCAAGAGGCGCCAG-----CAAGATGCGAGGGCC 2178  
Db 2407 CTAATGCGCTACAAGTGTGACGCGGCGCAGCGCCCGCGCAAGGATTCCTCGCGCT 2466  
Qy 2179 GTGACAACTGTACTCGCAGACCAACGCGCGCCAGCAGCAGCTCCCTTCTGGGGCGCAG 2238  
Db 2467 GTTAGCAGCTTGTCTCCAGACCAACAGCGCGCTTGGGCCCCACGCGCCCGCGCG 2526  
Qy 2239 GCGGGGCGCCCGCGCAGG 2257  
Db 2527 CCGCGCCCGGAGCGCGCGG 2545

RESULT 6

US-10-245-771-99  
; Sequence 99, Application US/10245771  
; Publication No. US20030068781A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3630R1C98  
; CURRENT APPLICATION NUMBER: US/10/245,771  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114

; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 99  
; LENGTH: 2855  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-245-771-99  
Query Match 19.8%; Score 621.8; DB 9; Length 2855;  
Best Local Similarity 61.3%; Pred. No. 1.3e-162;  
Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;  
Qy 507 GTTTCCTGTGTCAGCGCTGCCCAAGTACTGTCTGCCAGAAATCTGTCTGAGTCACT 566  
Db 798 CTCACACACACCCAGCCCATGTCCCGCGCTGCCGCTGCCAGACACAGTCTGCCCCCT 857  
Qy 567 GGGGACCTGTGCCCCCTCAAGGGCTGCTCTTTCTACCCCTGATATTGACCGCGGAC 626  
Db 858 AAGCGTGTGTGCCAGGGGAGGCTCTGTCTGTGCCACCTCTGCTGGACCGCGGCG 917  
Qy 627 AGTGAGCTGTGGCGCTGGCGGCAACTTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686  
Db 918 AGCGAGCTGTGGCTGGCAGACAACTTCATCGCTCCGTGCGCGCGCGGACCTGGCCAA 977  
Qy 687 CATGACGGGCTGTGTGACCTGACCTGTCCAGGAACACATCAGCCACATCAGCCCTT 746  
Db 978 CATGACGGCTGTGCTGATCTGAGCTGTCCGGAACACCATCCGCCACGTGGCTGCCG 1037  
Qy 747 TTCTTTCTGGACCTCGAGAGCTCCGCTCCCTGATCTTGACACAAATCGCTGCAAG 806  
Db 1038 CGCTTCGCCACCTGTGGGCGCTCGTGGCTGCTGCACTTGGATGCAACCGCTGACCTC 1097  
Qy 807 CCTTGGGAGGACACCTCCGCGGCTGTCAACCTGACGACCTTATCGTGAACAA 866  
Db 1098 ACTGGCGAGGGCAGCTGCGCGGCTGTGCAACTTGGCCACCTCATCTCAGCAACA 1157  
Qy 867 CCAGCTGGGCGGATCGCAGATGAGCTTTTGAGAGCTTCTGTGTGACATTTGAGAGATCT 926  
Db 1158 CCAGCTGGCAGCTGTGGCGGCGCGCTGTGATGATTGTGCCGAGACATCGGAGACCT 1217  
Qy 927 GGACCTCTCTACAAACCTTCCATGGCTGCGGTGGGACTCCGTGCGACGATGGTCAA 986  
Db 1218 CGACCTCTCTACAAACCTTCCATGGCTGCGGTGGGAGGCTTGGGCGGCTGGGCAA 1277  
Qy 987 CCTCACCAGCTGAGCTGGACCAACCTGCTGGATCACATCGCCGAGGCGACCTTTGC 1046  
Db 1278 CGTCAACAGTTGGGCTTCGACCAACCTGCTGCTTCTGTGCCC--GGGCTTTTTC 1334  
Qy 1047 AGACTGAGAACTGGCCCGCTCGATCTCACTTCAATCGGCTGCGAGAGCTGCCCC 1106  
Db 1335 CGGCTGCAAGCTGGCGCGCTGGACATGACCTTCAACCGCTGACCAATCCAC 1394  
Qy 1107 TGATCCCATTTTGGCGGCTCCAGCTTCCGCTTTGACAGCCACACCTTTTGGCCACC 1166  
Db 1395 CGACCACTTCTTCCCGCTGCCCCCTGCTCGCCAGGCGCGGCTCGCCGCTCTGCTGC 1454



Tue Jun 10 08:57:37 2003

1167 CTGTGCTTTAGTTTGGGGTAACCACTTCACTGCAATTTGAGCTTCTCTGGCTGG 1226  
1455 CCTGTGCTGGCTTTTGGCGGAACCCCTGCTGCACTGCAACTGCGAGCTGTGTGGCTGG 1514  
1227 GAGGCTCGAGCGGAGAGTACCTTGAAACCTGTGGCTCCCGAGGGGCGCTCAAGGGTGG 1286  
1515 TGCCCTGGCGGGAGGAGCACTCGAGGCTTGGCGTCCCGTCCCGACCTGCTCTGGGCGGCG 1574  
1287 CTACTTCTGGCATGTGCTGAGAGAGTGTGTGCGAGCGGCTCTCATCACCCAGCA 1346  
1575 CTACTTCTGGCGGTGGCGAGGAGTGTGTGCGAGCGGCTCTCATCACCCAGCA 1634  
1347 CACACACAAGTTGCTGCTTCTGAGGCGCAGCGGCACTCAAGTGCAGCAAGCATTTGG 1406  
1635 CTCACCACTCTGCTGTGCGCGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1694  
1407 GGACCCCGAGCCCTTATCCACTGGGTAGCCCGGATGACCGCTGTGAGGAATCTCTC 1466  
1695 GGACCCAGAGCCCGTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1754  
1467 AAGACCGCTGTATGACAACTGCACTGCACTGCACTTTCATCACCACTCTCAGGACAG 1526  
1755 CCGTGGCGCGCTTCCCAATGGAGCTGTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1814  
1527 TGGTGCCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1586  
1815 TGGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1874  
1587 CTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1644  
1875 GACTGTGGTCCCCCAGCCTCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1934  
1645 TCCGCTCTCAGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704  
1935 GGACGGGATCTGATGCTCTACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1994  
1705 GCGGAGAGCTTCCAAAGCCCGGAGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1764  
1995 CGACACTGGGCCCCCTA-----CGACCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046  
1765 ACCTGGCCCTGCTCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1824  
2047 ACAGTGTCTTGTTCAGTGGCGGATCAGCGGCTATCCCGGCACTCCGATGTACAG 2106  
1825 CTGAGTACAAGTGTCTGAGATGAGTATGATTTACAGATGATCCAGCTTCCAAAC 1884  
2107 ATCCAGTACAAGTGTCTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2166  
1885 AAGGCTTGTGCTCAACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1944  
2167 CGTGTGCT 2226  
1945 ATGTGGATGACACAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2004  
2227 GTGATGAGGAGCGCCAGCGGCTTCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2286  
2005 TTCAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2064  
2287 TCCACCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2346  
2065 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2124  
2347 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2406  
2125 ATGCT 2178  
2407 CTAATGCT 2466  
2179 GTGAGCAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2238  
2467 GTTACAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2526  
2239 GCGGGGCGCGCGCAGG 2257

Db 2527 CCGGCCCCGAGCCCGCG 2545  
RESULT 7  
US-10-245-851-99  
; Sequence 99, Application US/10245851  
; Publication No. US20030068792A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Baton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3630R1C93  
; CURRENT APPLICATION NUMBER: US/10/245,851  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 99  
; LENGTH: 2855  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-245-851-99  
Query Match 19.8%; Score 621.8; DB 9; Length 2855;  
Best Local Similarity 61.3%; Pred. No. 1.3e-162;  
Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;  
QY 507 GTTTGCGGTGGTGCAGCGCTGCCCCCAAGTACTGTGCTGCCAGAAATCTGCTGAGTCACT 566  
Db 798 GTCAGGCACACCCAGGCCCATGTCCCGCGCTGCCGCTGCCAGACACAGTCTGCTGCCCT 857  
QY 567 GGGACCTGTGCTCCCTCCAAAGGGGCTGCTTTTGTACCCCTGATATTGACCGGGGAC 626  
Db 858 AAGCGTGTGTGCCAGGGGAGCGCTCTGTTGTCGCCACCTCGCTGGACCGCGGGC 917  
QY 627 AGTGAGCTCGGCTGGGGGCAACTTCATCATCAGCCGCGGAGACTTTGCCAA 686  
Db 918 AGCCAGCTCGGCTGGCAGCAACTTCATCGCTTCCGCGCGCGGAGCTGGCCAA 977  
QY 687 CATCAGCGGCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746  
Db 978 CATGACAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1037

QY 747 TTCCTTTCTGACCTCGAGAGCCTCCGCTCCCTGCTGATTTGACAGCAATCGCTCCCAAG 806  
 Db 1038 GGCCTTCGCGACCTCGCGGCGCTCGTGCCCTGCTGACCTGGATGCAACCGGCTGACCTC 1097  
 QY 807 CTTGGGAGGACACCCCTCGGGGCGCTGGTCAACCTGAGCAGCACTTATCGTGAACAA 866  
 Db 1098 ACTGGCGAGGGCCAGCTGCGGGGCGCTGGTCAACCTGGGCAACCTCACTCTCAGCAAA 1157  
 QY 867 CCAGCTGGCGGCGCATCGAGATGAGGCTTTTGGAGACTTCTGCTGACATGGAGGATCT 926  
 Db 1158 CCAGCTGGCAGGCTGGCGGCGCGCTCGATGATGATGTCGCGAGACACTGGAGGACT 1217  
 QY 927 GGACTCTCTCAACAACTCCATGGCTGCGCTGGGACTCCGCTGGCAAGCATGGTCAA 986  
 Db 1218 CGACTCTCTCAACAACTCGAGCAGCTGCGCTGGGAGGCGCTGGGCGCGCTGGGCAA 1277  
 QY 987 CCTCACAGCTGAGCTGGACCAACACTGCTGATCATCGCGGAGGCACTTTGC 1046  
 Db 1278 GGTCAACAGTTGGGCGCTCGACCAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1334  
 QY 1047 AGACTGCGAGAACTGGGCGCGCTGGATCTCACTCCCAATCGGCTGCAGAACTGCGCCC 1106  
 Db 1335 CGGCTGCAAGCTGGCGGCGCTGGACATGACTTCCAACTCCAACTGACCAAACTCCACC 1394  
 QY 1107 TGATCCCATTTTGGCGGCTCCAGGCTTGGCTTTGACAGCACACCTTTTGGCCCAACC 1166  
 Db 1395 CGACCACTCTTCTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1454  
 QY 1167 TTTGCTCTTTAGTTTGGGGTAAACCACTTCACTGCAATGCTGCTGCTGCTGCTGCTG 1226  
 Db 1455 CTTGCTGCTGCTGCTTGGCGGGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1514  
 QY 1227 GAGGCTCGAGCGGAGCTGACTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1286  
 Db 1515 TCGCTTGGCGGGAGGAGCACTCGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1574  
 QY 1287 CTACTTCTGGCATGTGCGTGGAGGAGTTTGTGCGAGCGGCTTCTCATCACCAGCA 1346  
 Db 1575 CTACTTCTGGCGGTGGCGAGGAGGAGTTTGTGCGAGCGGCTGCTGCTGCTGCTGCTG 1634  
 QY 1347 CACACAAAGTTGCTGTTGCTGGAGGCGCAGCGGCGCACTCAAGTGCMAAGCACTGG 1406  
 Db 1635 CTCACCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1694  
 QY 1407 GGACCCAGCCCTTATCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1466  
 Db 1695 GGACCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1754  
 QY 1467 AAGGACCGCTGTATGCAATGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1526  
 Db 1755 CCGTGGCGGCGCTTCCCAATGGGAGCGCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1814  
 QY 1527 TGGTGGCTTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1586  
 Db 1815 TGGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1874  
 QY 1587 CTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644  
 Db 1875 GACTGTGGTCCCCCACCACCTCTCTAGCTAGCCAAACGACCACTGCTGCTGCTGCTGCTG 1934  
 QY 1645 TCCCGCTCTCAGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1704  
 Db 1935 GGACGGGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1994  
 QY 1705 GCGGAGAGCTTCCCAAAAGCCCCCGGAAAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764  
 Db 1995 CGACACTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2046  
 QY 1765 ACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824  
 Db 2047 ACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2106

QY 1825 CTGAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGATGATCCAGCCTCCAAC 1884  
 Db 2107 ATCCAGTACAAAGCTCGGCTGATGACATCTCTGCTACAGATGATCCCGCGGAGAGC 2166  
 QY 1885 AAGGCTTCTGCTGCTCAACAACTGCTGCTCAGGACTGCTGCTGCTGCTGCTGCTGCTG 1944  
 Db 2167 CGCTGCTTCTGCTGACGAGCTTGGCTGAGCGGACCTAGATCTGCTGCTGCTGCTGCTG 2226  
 QY 1945 ATGTGGGATGACAGACGACGACACTCAGGCGCACCAACATCGTGGGCTGCGCCCAAGTTC 2004  
 Db 2227 GTGTATGAGGACAGCGCCACGCGGCTCAGCGCCACGCGGCTGCTGCGGCTGCGCCGCTTC 2286  
 QY 2005 TTACCAAGGCTGACTACCGGAGTCCAGTCCATGACAGCCAGATTTCTGGGCGGCGACC 2064  
 Db 2287 TCCACCGAATCTGCTGCTGCGGCTGCGGCGCATGCGGCGCGCTGCTGCTGCTGCGGCGCAG 2346  
 QY 2065 ATGATCTCTGCTGCTGCGGCGCATCATGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTG 2124  
 Db 2347 ATGATCATGCTGCTGCGGCGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2406  
 QY 2125 ATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2178  
 Db 2407 CTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2466  
 QY 2179 GTGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2238  
 Db 2467 GTTAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2526  
 QY 2239 GCGGCGGCG 2257  
 Db 2527 CCG 2545

RESULT 8

US-10-245-883-99  
 ; Sequence 99, Application US/10245883  
 ; Publication No. US20030068783A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stephan, Jean-Philippe  
 ; APPLICANT: Watanabe, Colin  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; APPLICANT: Fong, Sherman  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P363081C70  
 ; CURRENT APPLICATION NUMBER: US/10/245,883  
 ; CURRENT FILING DATE: 2002-09-16  
 ; PRIOR APPLICATION NUMBER: 10/197942  
 ; PRIOR FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/059114  
 ; PRIOR FILING DATE: 1997-09-17  
 ; PRIOR APPLICATION NUMBER: 60/063046  
 ; PRIOR FILING DATE: 1997-10-24  
 ; PRIOR APPLICATION NUMBER: 60/065027  
 ; PRIOR FILING DATE: 1997-11-10  
 ; PRIOR APPLICATION NUMBER: 60/079689  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/086478  
 ; PRIOR FILING DATE: 1998-05-22  
 ; PRIOR APPLICATION NUMBER: 60/087607  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24

Tue Jun 10 08:57:37 2003

```

; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
; LENGTH: 2855
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-883-99

Query Match          19.8%; Score 621.8; DB 9; Length 2855;
Best Local Similarity 61.3%; Pred. No. 1.3e-162;
Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

QY 507 GTTTCGCTGTCGAGCGCTGCCCAAGTACTGTGTGCCAGAACTCTGTCTGAGTCACT 566
DB 798 GTACGACACACCCAGCCCATGTCCCGCGCTGCCGTGCCAGACACAGTCCGCTCCCT 857
QY 567 GGGGACCTGTGCCCCCTCCAAAGGGGCTGCTCTTTGACCCCTGATATTGACCGGGGAC 626
DB 858 AAGGTGCTGTGCCAGGGGACGCTCTGTTGTCGACACCTCGCTGGACCCGCGGC 917
QY 627 AGTGGAGCTGGCTGGGGGGAATTCATCATCCATCAGCGGAGGACTTTGCCAA 686
DB 918 AGCGAGCTGGCTGGGAGACAACTTCATGCTCCGTGCGGCGCGGAGACTGGCCAA 977
QY 687 CATGACGGGCTGTGTGACCTGACCTGTCTCCAGGAAACCATCAGCCATCCAGCCCTT 746
DB 978 CATGACAGGCTGTGATCTGAGCTGTCTGCGGAAACCATCCGCGACGTGGCTGCGG 1037
QY 747 TTCCTTTCTGACCTCGAGAGCTTCGCTCCCTGTCATCTGACGAAATCGGCTGCCAAG 806
DB 1038 CGCCTTCCGCGACCTGCGGCGCTGCGTGCCTGCTGACCTGATGGGAAACCGGCTGACCTC 1097
QY 807 CTTTGGGGAGACACCTCGGGGCTGTGTCAACTGTCAGACCTTATGCTGAACAA 866
DB 1098 ACTGGGAGGGGACAGTGGCGGCTGTGTCAACTTGGCCACCTCATCTCAGCAACAA 1157
QY 867 CCAGCTGGGGGACATCGAGATAGGCTTTTGGAGCTTCTGTGACATTGGAGATCT 926
DB 1158 CCAGCTGGAGCGTGGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1217
QY 927 GGACCTCTCTACACACCTCCATGCGCTCGGAGCTTCCGTGGAGCGCATGATGTC 986
DB 1218 CGACCTCTCTACACACCTCCATGCGCTCGGAGCTTCCGTGGAGCGCATGATGTC 1277
QY 987 CTTCCACACAGTGGAGCTGGACCAACCTGCTGATCACAATCGCGAGGGACCTTTGC 1046
DB 1278 CGTCAACAGTGGGCGCTCGACCAACCTGCTGATGATGATGATGATGATGATGATGATGAT 1334
QY 1047 AGACCTGCAAGAACTGGCGCGCTGATCTCACCTCCAACTCGGTCGAGAGCTGCGCC 1106
DB 1335 CGGCTTGCACAGCTGGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1394
QY 1107 TGATCCCACTTTGCGCGCTTCCAGGCTTTCGAGCTTTCGAGCAACACCTTTGCGCCACC 1166
DB 1395 CGACCACTTCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1454
QY 1167 CTTGCTCTTTAGTTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTCTGCTGCG 1226
DB 1455 CTTGCTCTTTAGTTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTCTGCTGCG 1514
QY 1227 GAGGCTCGAGCGGAGATGACCTGGAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1286
DB 1515 TCGCCCTGGCGGGAGAGACCTCGAGGCTTGGCGTCCCACTGCTGCTGCTGCTGCTGCTGCTG 1574
QY 1287 CTACTTCTGGCATGTGCTGAGGAGGATTTGTCGAGCGGCTCTCTATCACCAGCA 1346
DB 1575 CTACTTCTGGCGGTGGCGAGGAGATTTGTCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1634
QY 1347 CACACACAGTTGCTGTTGAGGCGGAGGCGGACACTCAAGTGCAGAGCCATTTGG 1406
DB 1635 CTCACCACTCTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1694

```

```

QY 1407 GGACCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCCGCTGTGGTGGAACTCTCTC 1466
DB 1695 GGACCCAGAGCCCCGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1754
QY 1467 AAGGACCGCTGTATGACAAATGGACCTGAGCATTTTCATCACCACATCTCAGGACAG 1526
DB 1755 CCGTCCCGCGCTTCCCAATGGGACGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1814
QY 1527 TGGTGCCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1586
DB 1815 TGGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1874
QY 1587 CTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
DB 1875 GATGTGGGTCCCCCACCCTCTCAGCTAGCAACAGCAGCTGTGACCCCGCGG 1934
QY 1645 TCCCGCTCTCAGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1704
DB 1935 GGAAGGGGATCTGATGCTCTCACCACCTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1994
QY 1705 GCGGAGAGCTCCCAAAAGCCCCCGGAAAGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1764
DB 1995 CGACACTGGGCCCCCTA-----CCGACCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2046
QY 1765 ACCTCGGCTTGGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824
DB 2047 ACAGCTCTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2106
QY 1825 CTGACGTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1884
DB 2107 ATCCAGTACACAGCTCGGCTGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2166
QY 1885 AAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
DB 2167 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2226
QY 1945 ATGTGGGATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2004
DB 2227 GTGTATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2286
QY 2005 TTCAACAAAGGCTGACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2064
DB 2287 TCACCGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2346
QY 2065 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2124
DB 2347 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2406
QY 2125 ATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2178
DB 2407 CTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2466
QY 2179 GTGAGCAATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2238
DB 2467 GTTAGCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2526
QY 2239 GCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2257
DB 2527 CCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2545

```

RESULT 9  
 US-10-237-535-99  
 ; Sequence 99, Application US/10237535  
 ; Publication No. US20030073188A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney/Austin

APPLICANT: Smith,Victoria  
APPLICANT: Stephan,Jean-Phillippe  
APPLICANT: Watanabe,Colin  
APPLICANT: Wood,William  
APPLICANT: Zhang,Zemin  
APPLICANT: Fong,Sherman  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P36301C3  
CURRENT APPLICATION NUMBER: US/10/237,535  
CURRENT FILING DATE: 2002-09-06  
PRIOR APPLICATION NUMBER: 10/197942  
PRIOR FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/059114  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/063046  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/065027  
PRIOR FILING DATE: 1997-11-10  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/086478  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090689  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/091358  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/099803  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/106932  
PRIOR FILING DATE: 1998-11-03  
PRIOR APPLICATION NUMBER: 60/115554  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/119342  
PRIOR FILING DATE: 1999-02-09  
PRIOR APPLICATION NUMBER: 60/123957  
PRIOR FILING DATE: 1999-03-12  
PRIOR APPLICATION NUMBER: 60/123972  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: 60/127372  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/131271  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/133459  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/135725  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/135729  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/135750  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/138385  
PRIOR FILING DATE: 1999-06-09  
PRIOR APPLICATION NUMBER: 60/140653  
PRIOR FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: 60/141037  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: 60/144732  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/144758  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/144790  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/145228  
PRIOR FILING DATE: 1999-07-20

PRIOR APPLICATION NUMBER: 60/145698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: 60/146222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: 60/146843  
PRIOR FILING DATE: 1999-08-03  
PRIOR APPLICATION NUMBER: 60/148188  
PRIOR FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/148513  
PRIOR FILING DATE: 1999-08-12  
PRIOR APPLICATION NUMBER: 60/149327  
PRIOR FILING DATE: 1999-08-17  
PRIOR APPLICATION NUMBER: 60/149395  
PRIOR FILING DATE: 1999-08-17  
PRIOR APPLICATION NUMBER: 60/150114  
PRIOR FILING DATE: 1999-08-20  
PRIOR APPLICATION NUMBER: 60/151700  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: 60/151734  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: 60/162506  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 60/170262  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 60/177118  
PRIOR FILING DATE: 2000-01-20  
PRIOR APPLICATION NUMBER: 60/179851  
PRIOR FILING DATE: 2000-02-02  
PRIOR APPLICATION NUMBER: 60/180921  
PRIOR FILING DATE: 2000-02-08  
PRIOR APPLICATION NUMBER: 60/187202  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 60/198587  
PRIOR FILING DATE: 2000-04-18  
PRIOR APPLICATION NUMBER: 60/199614  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 60/206330  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/206368  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/209832  
PRIOR FILING DATE: 2000-06-05  
PRIOR APPLICATION NUMBER: 60/218371  
PRIOR FILING DATE: 2000-07-13  
PRIOR APPLICATION NUMBER: 60/222695  
PRIOR FILING DATE: 2000-08-02  
PRIOR APPLICATION NUMBER: 60/229896  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/230621  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/232887  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: 60/235147  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 60/261878  
PRIOR FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: 60/261910  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: 60/261939  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: 60/262150  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: 60/264395  
PRIOR FILING DATE: 2001-01-25  
PRIOR APPLICATION NUMBER: 60/266421  
PRIOR FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: 60/267623  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/274399  
PRIOR FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: 60/280982  
PRIOR FILING DATE: 2001-04-03  
PRIOR APPLICATION NUMBER: 60/282129



QY 1945 ATGTGGATGACACAGCCACGACACTCAGCGCCACCAACATCGTGGGCTGCCCGCAGTTC 2004  
Db 2227 GTGTATAGGACAGCGCCACCGGGCTCAGCGCCAGCGGCTGTGGGCTGCCCGCCTTC 2286  
QY 2005 TTCACCAAGGCTGACTACCCGAGTGGCCAGTCCATGACAGCCAGATTTCTGGGGCGGACC 2064  
Db 2287 TCCACCGAACCTGCGCTGGGCCATGGGGGGCGCCGACGCTCCCTTCTGGGGCGGACG 2346  
QY 2065 ATGATCTCTGGTATCGGGGGCATCATCGTGGCCAGCTGCTGGTCTTCATCGTCACTCCTC 2124  
Db 2347 ATGATCATCGCGCTGGCGGCGTCTATCGTAGCTCGGTACTGGTCTTCATCTTCTGTGCTG 2406  
QY 2125 ATGTGGCGGTACAGGTCTGCAACACGAGGCCCCCAG-----CAAGATGCGACGGGCC 2178  
Db 2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCGCCCCCGGCAAGGCGCAAGATTCCTCGCGCT 2466  
QY 2179 GTGACCAATGTGTACTCGCAGACCAACGCGCGCCAGCCACTCCCAAGCAGCGCACCA 2238  
Db 2467 GTTACAGCGTTTGTCTCCAGACCAACGGCGCCCTTGGGCCCCACGCGCCCGCCCG 2526  
QY 2239 CCGCGGGCCCCCGCCGAGG 2257  
Db 2527 CCGCGCCCGGAGCCCGCGG 2545

RESULT 10

US-10-238-183-99  
; Sequence 99, Application US/10238183  
; Publication No. US20030073189A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watambe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3630R1C11  
; CURRENT APPLICATION NUMBER: US/10/238,183  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/091358  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/099803  
; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: 60/106932  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: 60/115554  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/119342  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: 60/123957  
; PRIOR FILING DATE: 1999-03-12  
; PRIOR APPLICATION NUMBER: 60/123972  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: 60/127372  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/131271  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/133459  
; PRIOR FILING DATE: 1999-05-11  
; PRIOR APPLICATION NUMBER: 60/135725  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 60/135729  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 60/135750  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 60/138385  
; PRIOR FILING DATE: 1999-06-09  
; PRIOR APPLICATION NUMBER: 60/140653  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/141037  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 60/144732  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/144758  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/144790  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145228  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: 60/146222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: 60/146843  
; PRIOR FILING DATE: 1999-08-03  
; PRIOR APPLICATION NUMBER: 60/148188  
; PRIOR FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/148513  
; PRIOR FILING DATE: 1999-08-12  
; PRIOR APPLICATION NUMBER: 60/149327  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/149395  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/150114  
; PRIOR FILING DATE: 1999-08-20  
; PRIOR APPLICATION NUMBER: 60/151700  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/151734  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/162506  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: 60/170262  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 60/177118  
; PRIOR FILING DATE: 2000-01-20  
; PRIOR APPLICATION NUMBER: 60/179851  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 60/180921  
; PRIOR FILING DATE: 2000-02-08  
; PRIOR APPLICATION NUMBER: 60/187202  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: 60/198587  
; PRIOR FILING DATE: 2000-04-18  
; PRIOR APPLICATION NUMBER: 60/199614  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 60/206330







987	QY	CCTCCACAGCTGAGCGCTGGACCAACCTGCTGGATCATGCGGAGGGCACTTTGC	1046
1278	DB	CGTCAACACGTTGGGGCTCGACCACAACCTGCTGGCTCTGTGGCC--GGCGCTTTTTC	1334
1047	QY	AGACCTGCAGAAACTGGCCCGCTGGATCTCACTCCAATCGGCTGCAGAACTGCCCCC	1106
1335	DB	CCGCTGCAACAGTGGCCCGCTGACATGACCTCAACCCCTGACCACAATCCACC	1394
1107	QY	TGATCCCATTTTGGCCGCTCCAGGCTTGGCTTTGACAGCCACACCTTTTGCCCCACC	1166
1395	DB	CGACCCACTCTTCTCCCGCTGCCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCTCTGC	1454
1167	QY	CTTGCTCTTTAGTTTGGGGGTAAACCACTCACTGCAATTGTAGACTTCTCTGGCTGG	1226
1455	DB	CCTGGTCTGGCCCTTTGGCGGGAACCCCTGCACTGCAATGCGAGCTGGTGTGGCTGG	1514
1227	QY	GAGGCTCGAGCGGACGATCACTTGGAAACCTGTGGCTCCCAAGGGGGCTCAAGGGTGG	1286
1515	DB	TCGCTGGCGGGAGGACGACCTCGAGGCTTGGCGTCCCACTGCTCTGGGGCGCG	1574
1287	QY	CTACTTCTGGCATGTGGTGGAGGAGTTGTGTGGAGCCGCTCTCATCACCAGCA	1346
1575	DB	CTACTTCTGGCGGTGGCGGAGGAGTTGTGTGCGAGCCCGCTGGTGAATCACC	1634
1347	QY	CACACAAAGTTGCTGTTCTGGAGGCGCAGGCGGCCACACTCAAGTGCAAAGCCATTGG	1406
1635	DB	CTACACACTTGGCTGTGCCCGCAGGTGGCGGCTGCCCTGCGTGC CGGSCAGTGGG	1694
1407	QY	GGACCCACGCCCTTATPCCACTGGGTAGCCCGGATGACGGCTGGTAGGGAATCTCTC	1466
1695	DB	GGACCCAGAGCCCCGTGTGCTTGGGTGTCACCCCAGGCGCGGTGTAGGCAATCAAG	1754
1467	QY	AAGGACCGCTGCTATGACATGACACCCCTGGACATCTTCATCACCACATCTCAGACAG	1526
1755	DB	CCGTGCCCGCGCTTCCCAATGGGACGCTGGAGCTGCTGTTACCAGCCGGGTATGG	1814
1527	QY	TGGTGCCCTTCACTGCTTGTGCAATGCTTGGCGGAGGCCACGGCCATGTTGAGGT	1586
1815	DB	TGGCATCTTCACTGATTGGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT	1874
1587	QY	CTCCATCGTCCAGCTGCCACACTTCAGCAACAGCACCAGCCGGA--CTGCACCCCCCAAG	1644
1875	DB	GACTGTGGGTCCCCACCACTCTCAGCTAGCCAAAGCACCAGCTGTGACCCCCCGCG	1934
1645	QY	TCCCGCCTCTCAGACATCACTTGGTCTCAGCAAGACGACGCGGAGGTGGAGGCACTGGG	1704
1935	DB	GGACGGGATCTGATGCTCTCACCCACCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC	1994
1705	QY	GGCGGAGAGCTTCCAAAGCCCCCGGAAACGGGCTGTGTTGTCTGAAGTGCACCACC	1764
1995	DB	CGACACTGGGCCCTTA-----CCGACCGTGGCGTCCAGGTGACTGACACGCGGGCC	2046
1765	QY	ACCTCGGCCCTGGTCAAGTGCTCTCAGCAAGTCAGCACACCCCGGTGAAGATCTACAG	1824
2047	DB	ACAGTGTCTTGTCCAGTGGCCGATCAGCGGCTTATCCCGGCAATCCGATGTATACAG	2106
1825	QY	CTGCAGTCAAACTGCTCTGACATGAGGTACTGATTTTACAGGATGATCCCAAGCTCAAC	1884
2107	DB	ATCCAGTACAAAGCTCGGCTGATGACATCCTCGTCTACAGATGATCCCGCGGAGAGC	2166
1885	QY	AAGGCTTCGTGGTCAACAACTGGTGTTCAGGACTGGCTAGACTTGTGTGTCTGGCC	1944
2167	DB	CGCTGCTTCTGCTGACGGACCTGGCGTCAGGCCGAGCCTACGATCTGTGGTGTCTGCC	2226
1945	QY	ATGTGGATGACACAGCCACGACATCAAGCGCCACCAACATCGTGGGTGGCGCCAGTTC	2004
2227	DB	GTGTATGAGCAGCGCCACGGGCTCAGGCCACCGCGCTGTGGCTGCGCCGCTTC	2286
2005	QY	TTACCAAAGGTGACTACCGCAGTGCCAGTCCATGCAACGCCAGATTCTGGGGGACAC	2064
2287	DB	TCCACCGAACTGCTGCTGGGCAATCGGGGGCGCGCACGTCCCTTCTTGGGGCAGC	2346

Qy	2065	ATGATCTGTGTACATCGGGGGCATCATCGTGGCCACCGTGTCTGCTTTCATCGTCATCCTC	2124
Db	2347	ATGATCATCGCTGGCGCGGTATCGTAGCCTCGGTACTGTGTCTTTCATCTTCTGTGCTG	2406
Qy	2125	ATGGTGGCTACAAAGTCTGCAACCAAGAGGCCCCCAG-----CAAGATGGCAGCGGCC	2178
Db	2407	CTAATGGCTTACAAAGGTGTCAAGGGGGCCAGCCCCCGCAAGCAAGATTCCGCGCCT	2466
Qy	2179	GTGAGCAATGTACTCGCAGACCAACGGCGCCAGCCACCGCTTCAAGCAGGCGACCA	2338
Db	2467	GTTAGCAGCGTTGTTCTCCAGACCAACGGCGCGCTGGGGCCCCCAGCCCGCGCCCG	2526
Qy	2239	CGCGGGCCCCCGCGCAGG	2257
Db	2527	CCGCGCCGAGCGCCGCG	2545

RESULT 12

US-10-238-370-99

; Sequence 99, Application US/10238370

; Publication No. US20030073191A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin

; APPLICANT: Eaton, Dan

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Goddard, Audrey

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin

; APPLICANT: Smith, Victoria

; APPLICANT: Stephan, Jean-Phillippe

; APPLICANT: Watanabe, Colin

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; APPLICANT: Fong, Sherman

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3630R1C10

; CURRENT APPLICATION NUMBER: US/10/238,370

; CURRENT FILING DATE: 2002-09-09

; PRIOR APPLICATION NUMBER: 10/197942

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/059114

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/063046

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/065027

; PRIOR FILING DATE: 1997-11-10

; PRIOR APPLICATION NUMBER: 60/079689

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/086478

; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-08-18

; PRIOR APPLICATION NUMBER: 60/090557

; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090689

; PRIOR FILING DATE: 1998-06-25

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 116

; SEQ ID NO 99

; LENGTH: 2855

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-238-170-99

Query Match	19.8%;	Score 621.8;	DB 9;	Length 2855;
Best Local Similarity	61.3%;	Pred. NO. 1.3e-162;		
Matches 1078;	Conservative	0;	Mismatches 662;	Indels 19; Gaps 4
507	GTWTCGGTGTGCAGCCCTGCCCCAAAGTACTGTGTCTGCCAGAAATCTGTCTGAGTCACCT			566

Db 798 GTACGCCACACCCAGGCCATGTCCCGCGCTGCCGCTGCCAGACACACAGTCGCTGCCCT 857  
 Qy 567 GGGGACCTGTGCTCCCTCAAGGGGCTGCTCTTTGTACCCCTGTATATTGACCGCGGAC 626  
 Db 858 AAGCGTGTGCTGCCAGGGGAGGCTCTCTGTTGCTGCGCACCTCGCTGGACCGCGGGC 917  
 Qy 627 AGTGAGCTGCGCTGTGGCGGCAATTCATCATTCACATCAGCAGCGCAGGACCTTTGGCNA 686  
 Db 918 AGCGAGCTGCGCTGTGGAGAACATTCATCGCTCTCGTGGCGCGCGACCTGGCCAA 977  
 Qy 687 CATGACGGGCTGTGGACCTGACCTGTCAGAGAACACCATCAGCCACATCCAGCCCTT 746  
 Db 978 CATGACGGCTGTGCTGATCTGAGCTGTGCGGGAACACCATCCCGACCTGGCTGCGCG 1037  
 Qy 747 TTCTTTTCTGACCTCGAGAGCTCCGCTCCCTGCTATCTTGACAGCAATCGGCTGCCAAG 806  
 Db 1038 CGCTTTGCGGACCTGCGGGCTCGTGTGCTGACCTGATGCAACCGCTGACCTC 1097  
 Qy 807 CTTGGGAGGACACCTCGGGGCTGTGTAACCTGACAGACCTTATCGTGAAACAA 866  
 Db 1098 ACTGGGCGAGGGCCAGCTGCGGGCTGTGTCAACTTTGGCCACCTCATCTCAGCAACAA 1157  
 Qy 867 CCAGCTGGGCGCATCGAGATGAGCTTTTTCAGGACTTCTGCTGACATTTGGAGATCT 926  
 Db 1158 CCAGCTGGAGCGCTGGCGGCGGCTGTGATGATTGTGCGGAGACATCGAGACCT 1217  
 Qy 927 GGACCTCTCTTAAACAACTCCATGCGCTGCGTGGGACTCCGTGCGAGCGATGTCAA 986  
 Db 1218 CGACCTCTCTTAAACAACTCGAGAGCTGCTTGGGAGGCTTGGGCGCTGGGCA 1277  
 Qy 987 CTTCCACAGCTGAGCTGGACCAAACTGTGTGATCACAATCGCGAGGCACTTTGC 1046  
 Db 1278 CGTCAACAGCTGGGCTCGACCAAACTGTGTGCTTCTGTGCTTTC 1334  
 Qy 1047 AGACCTGAGAACTGGCGGCTGGATCTCACTCCATCGCTGCGTGCAGAACTGCGCCC 1106  
 Db 1335 CGGCTGCAAGCTGGCGGCTGGATGATGACCTTCAACCGCTGACCAAACTCCACC 1394  
 Qy 1107 TGATCCCATTTTGGCGCTCCAGGCTTCCGCTTTGACAGCAACACCTTTTGGCCAC 1166  
 Db 1395 CGACCACTTCTTCCGCTGCGCTGCTGCGGAGGCTTGGCGCTGCTGCTGCTGCTG 1454  
 Qy 1167 CTTGCTCTTTTGGGGTAACTTCACTTCAATTTGAGCTTCTTGTGCTGCG 1226  
 Db 1455 CTTGCTGCTGCTTGGCGGAACTTCCCTGCACTGCACTGCACTGCTGCTGCTGCTG 1514  
 Qy 1227 GAGGCTGAGGGAGGATGACTTGAACCTTGTGCTCCCGAGGGGCTCAAGGCTG 1286  
 Db 1515 TCGCTGGCGGGAGGAGCTCGAGGCTTGGCGCTTCCCACTGCTTGGCGGCG 1574  
 Qy 1287 CTACTTTGGCATGTGCTGAGGAGTGTGTCGAGCGGCTTCTCATCACCAGCA 1346  
 Db 1575 CTACTTTGGCGGGGAGAGAGTGTGTCGAGCGGCTTGGCTGCTGCTGCTGCTGCTG 1634  
 Qy 1347 CACACAAAGTTGCTGTTTCTGGAGGCGAGCGGCACTCAAGTGCAAAAGCCATTGG 1406  
 Db 1635 CTCACCACTCTGCTGCTGCGGAGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1694  
 Qy 1407 GAGCCCGAGCCCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1466  
 Db 1695 GGACCCAGAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1754  
 Qy 1467 AAGGACCGCTGTATGAAATGGACCTGACATCTTTCATCACCACATCTCAGACAG 1526  
 Db 1755 CCGTGGCGGCTTCCCAATGGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1814  
 Qy 1527 TGGTGCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1586  
 Db 1815 TGGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1874  
 Qy 1587 CTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644  
 Db 1875 GACTGTGGTCCCGACCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1934

Qy 1645 TCCGCTCTTCCAGACATCACTGGTCTCCAGCAAGACACCGCGGGAGGTGGAGGCACTGG 1704  
 Db 1935 GGACGGGATCTGTATGTCTTCAACCCACCTCCGCTGCTGCTTCTGCTCAAGTGGC 1994  
 Qy 1705 GCGGAGAGCTTCCAAAGCCCCCGGACCGGGCTGTGCTTGTCTGTAAGTGAACACC 1764  
 Db 1995 CGACACTGGGCCCCCTA-----CCGACCGTGGCGTCCAGGTGACTGAGCAGCGGGC 2046  
 Qy 1765 ACCTCGGCTTGGTCAAGTGGTCTGTAGCAAGTCAAGCAGCCCGGTGAAGATGTACACAG 1824  
 Db 2047 ACAGTGTCTTGTCTAGTGGCGGATCAGCGGCTATCCCGGATCCGATGTACACAG 2106  
 Qy 1825 CTGCACTAACTGTCTGAGATGAGTGTGATTTACAGATGATCCAGCCTTCAAC 1884  
 Db 2107 ATCCAGTAAACAGCTCGCTGATGACATCTCTGCTACAGATGATCCCGCGGAGAGC 2166  
 Qy 1885 AAGGCTTGGTCAACAACTGTGTGAGGACTGCTAGACTTGTGTGCTGCTGCTG 1944  
 Db 2167 CGCTGCTTCTGCTGACGAGCTTGGCTGACGCGGACCTACGATCTGTGCTGCTGCTG 2226  
 Qy 1945 ATGCGGATGACACAGCAGCAGCTCAGGCGCACAACTCGTGGGCTGCGCCAGTTTC 2004  
 Db 2227 GTGATAGGACAGCGCCACGCGGCTCAGGCGCAGCGGCTGTGGCTGCGCCGCTTC 2286  
 Qy 2005 TTCAACAGGCTGACTACCCGAGTCCAGTCCAGTCCATGACAGCCAGATTTCTGGCGGCA 2064  
 Db 2287 TCCACCAAGCTGCTGCTGCGGCTGCGGCGGCGCGCAGCTTCTTCTGCGGCGCAG 2346  
 Qy 2065 ATGATCTTGTCTATCGGGGCTATCTGTGGCAGCTGCTGCTTCTATCTGCTATCTTC 2124  
 Db 2347 ATGATCATCGGCTGCGGCGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 2406  
 Qy 2125 ATGCTGCGCTCAAGGCTGCAACACAGGCGGCGGCGGCGGCGGCGGCGGCGG 2178  
 Db 2407 CTATGCTGCTGAGGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2466  
 Qy 2179 GTGACCAATGTGCTGCTGAGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2238  
 Db 2467 GTTACAGGCTTGTCTCCAGCAACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2526  
 Qy 2239 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2257  
 Db 2527 CCGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2545

RESULT 13

US-10-245-055-99  
 ; Sequence 99, Application US/10245055  
 ; Publication No. US20030073192A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stephan, Jean-Phillippe  
 ; APPLICANT: Watanabe, Colin  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; APPLICANT: Fong, Sherman  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P36301C88  
 ; CURRENT APPLICATION NUMBER: US/10/245,055  
 ; CURRENT FILING DATE: 2002-09-16  
 ; PRIOR APPLICATION NUMBER: 10/197942  
 ; PRIOR FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/059114  
 ; PRIOR FILING DATE: 1997-09-17  
 ; PRIOR APPLICATION NUMBER: 60/063046

```
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
; LENGTH: 2855
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-245-055-99

Query Match          19.8%; Score 621.8; DB 9; Length 2855;
Best Local Similarity 61.3%; Pred. No. 1.3e-162;
Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

QY 507 GTTGGCGGTGCGAGCGCTGCCCAAGTACTGTGTGCGCAGAACTCTGTGAGTCACT 566
DB 798 GTGAGCCACACCCAGGCCATGTCCCGCGCGTCCCGCTGCCAGACAGTCGCTGCCCT 857

QY 567 GGGGACCTGTGCCCTCCAAAGGGGTGCTTTGTAGCCCTGATTAATGACCGGGGAC 626
DB 858 AAGCGTGTGTGCCAGGGCAGGCTCTCTGTGTCGACCCCTGATTAATGACCGGGG 917

QY 627 AGTGGAGCTGCGCTGGCGGCGCAACTTCATCATCCACATCAGCGCGCAGGACTTTGCCAA 686
DB 918 AGCGAGCTGCGGTGCGAGACAACTTCATCGCTCCGTGCGCGCGCGGACCTGGCCAA 977

QY 687 CATGACGGGCTGTGTGAGCTGACCTGTCCAGGAACACCATCAGCCACATCAGGCCCTT 746
DB 978 CATGACAGGCTGTGTGATCTGAGCTGTGCGGAAACACCATCGGCAACGTGCTGCCGG 1037

QY 747 TTCCTTCTGAGCTCGAGAGCTCGCGCTCCCTGCACTTTGACAGCAATCGGCTGCCAAG 806
DB 1038 CGCCTTCGCGACCTCGGGCCCTGCTGCTGCTGCACTGATGGCAACCGGCTGACCTC 1097

QY 807 CTTTGGGAGGACACCTTCGCGGGCTGTGTGCACTGCGAGCCTTATCGTGAAACAA 866
DB 1098 ACTGGCGAGGGCCAGCTGGCGGCGCTGGTCAACTTGGCCACCTCATCTCAGCAACAA 1157

QY 867 CCAGCTGGCGGATCCAGATGAGGCTTTTGGAGCTTCTGTCGACATTTGAGGATCT 926
DB 1158 CCAGCTGGCGGCTGTGGCGGCGCGCTGATGATTTGTCGCGAGACATGAGGACCT 1217

QY 927 GGACTCTCTTACAAACCTTCCATGCGCTGCGGGAATCTCGTGCGAGCATGTGTCAA 986
DB 1218 CGACTCTCTTACAAACCTTCCAGAGCTGCGCTGGAGGCGCTGGGCGCTGGGCAA 1277

QY 987 CTTCCACAGCTGAGCTGGACCAACACTGTGTGGATCAATCGCCGAGGGCACCITTCG 1046
DB 1278 CGTCAACAGCTTGGGCTTCGACCAACAACTGTGTGGCTTCTGTGCCC---GGCGCTTTTC 1334

QY 1047 AGACTCGAGAACTGGCGGCTGATCTCACTTCAATCGGCTGCGAGAGCTGCCCCC 1106
DB 1335 CCGCTGTGCAAGCTGGCGGCTGGATGATGATCTCAACCGCTGACCAATCCAC 1394

QY 1107 TGATCCCATCTTTCGCGCTCCCAAGGCTTGGCTTTTGACAGCCACACCTTTGCCCA 1166
DB 1395 CGACCCACTCTTCTCCGCTGCGCTGCTGCGAGGCGCGGCGCTGCGCGCTCTGC 1454

QY 1167 CTTGCTCTTTAGTTTGGGGTAAACCACTTCACTGCAATGTGAGCTTCTCTGGCTGCG 1226
```

```
DB 1455 CCTGTGCTGCGCTTTGGCGGGAACCCCTTGCACTGCAACTGCGAGCTGTGTGGCTGCG 1514
QY 1227 GAGGCTGAGCGGACATGACCTGGAACCTGTGGCTCCCGAGGGGCGCTCAAGGCTCG 1286
DB 1515 TCGCTGCGCGGAGGACGACCTCGAGGCTCGCGCTGCCACCTCTCTGGGCGCGCG 1574
QY 1287 CTACTTCTGGCATGTGCGTGAAGAGATTGTGTGGAGCGCGCTCTCATCACCAGCA 1346
DB 1575 CTACTTCTGGCGGTGGCGAGGAGATTGTGTGGAGCGCGCTGTGTGATCTACCG 1634
QY 1347 CACACAAAGTTGCTGTTCTGGAGGCGGAGGCGGCGACACTCAAGTGCAAGCAATTGG 1406
DB 1635 CTCACCACTCTGCTGTGCCCGCAGGTGCGCGGCTGCGCTGCGCTGCCGGGAGTGG 1694
QY 1407 GGACCCAGCGCCCTTATCCACTGGTAGCCCCCGATGACCGCTGTGTAGGAACTCTC 1466
DB 1695 GGACCCAGAGCGCGCTGTGCTGGTGTCAACCCAGGCGGCGTGTAGGCAACTCAAG 1754
QY 1467 AAGGACCGCTGTCTATGACAAATGCGACCTGTGGACATCTTATCACCACATCTCAGGACAG 1526
DB 1755 CCGTGCCCGCGCTTCCCAATGGGAGCTGTGTGTCACCGAGCGCGGTGATGG 1814
QY 1527 TGGTGCCTTCACTGCAATTTGCTGCCAATGCTGCCGAGAGCGCACCGCATGTGTGGAGGT 1586
DB 1815 TGGCATCTTCACTGCAATTTGCGGCAATGCACTGTGGGAGGCGCACAGCTGTGTGGAGCT 1874
QY 1587 CTCCATCTGTCAGCTGCGCACACCTCAGCAACAGCACAGCCGCA--CTGACACCCCGCAAG 1644
DB 1875 GACTGTGGTCCCCACCACTCTCTCAGTAGCAACAGCACAGCTGTGACCCCGCG 1934
QY 1645 TCCGCGCTCTCAGACATCAGTGGTCCAGCAAGACCGGAGGCGGAGGCGAGGCGG 1704
DB 1935 GGACGGGATCTCTGATGCTCTCACCCACCTCCGCTGCTCTGCTTCTGCAAGGTGGC 1994
QY 1705 GCGGAGAGGCTCCCAAGAGCGGCTGTGTGTGCTGTGAAAGTCAACACC 1764
DB 1995 CGACACTGGGCGCCCTA-----CGACCGTGGCTTCAGGTGACTGAGCACGGGCGC 2046
QY 1765 ACCTCGGCGCTGTCAAGTGTGTGTGAGCAAGTCAAGCACCGCGGTTGAAGATGACAG 1824
DB 2047 ACAGCTGCTGTGTCCAGTGGCGGATCAGCGGCTATCCCGGGCATCCGATGTACAG 2106
QY 1825 CTGAGTACAACTCTCTGAGATGAGTACTGATTTACAGATGATCCAGCTTCAAC 1884
DB 2107 ATCCAGTACAACTGCTGCTGATGACATCTCTGCTTACAGATGATCCCGCGGAGAGC 2166
QY 1885 AAGGCTTCTGTGCTCAACAACTGTGTGAGGAGTGTGCTACGACTTGTGTGTGTGGCC 1944
DB 2167 CGCTGCTTCTGCTGACGACCTGCGGTGAGCGGCTACGATCTGTGCTGTGCTGCGC 2226
QY 1945 ATGTGGGATGACAGCCACGACACTCAGGCGCACCAACATGTTGGGCTGCGGCCAGTTC 2004
DB 2227 GTGTATAGGACAGCGCCACGCGGCTCAGCGGCTGTGGGCTGCGGCCGCTTC 2286
QY 2005 TTCACCAAGGCTGACTACCGCAGTCCAGTCCATGACAGCGAGATCTGGGCGGCGAC 2064
DB 2287 TCACCGAACTCTGCGGTGCGGCAATGGGGGCGCGCGACGCTCCCTTCTGTGGGCGCAG 2346
QY 2065 ATGATCTCTGCTCATCGGGGCGATCATCGTGGCGACGCTGTGCTTTCATGTCATCTC 2124
DB 2347 ATGATCATCGCTGGGCGGCTCATCGTAGCTCGTGTGCTTCTCATCTTCTGTGCTG 2406
QY 2125 ATGTGCTGCTTCAAGGTCTGCAACACAGAGGCGCCAG-----CAAGATGCGAGCGGC 2178
DB 2407 CTAATTCGCTTCAAGGTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2466
QY 2179 GTGAGCAATGTGTACTCGCAGACCAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2238
DB 2467 GTTAGCAGCTTGTCTCCAGACCAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2526
QY 2239 GCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2257
DB 2527 CCGCGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2545
```

```

RESULT 14
US-10-245-147-99
; Sequence 99, Application US/10245147
; Publication No. US20030073193A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C72
; CURRENT APPLICATION NUMBER: US/10/245,147
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
; LENGTH: 2855
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-147-99

Query Match          19.8%; Score 621.8; DB 9; Length 2855;
Best Local Similarity 61.3%; Pred. No. 1.3e-162;
Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

Qy 507 GTTTGCCGCTGCTCCAGCGCTGCCCCCAAGTACTGTGTCTGCCAGAACTCTGTGAGTCACT 566
Db 798 GTCAGCCACACCCAGGCCATGTCCCGCGCTGCCGCTGCCAGACAGTCTGCTGCCCT 857

Qy 567 GGGGACCCCTGTGCCCTTCCAAAGGGGTGCTCTTTGTATCCCCCTGATATTGACCGCGGAC 626
Db 858 AAGCGTGTGTGCCAGGGGAGCGCTCTCTGTTGTCACCCCTCGCTGGACCGCGCGG 917

Qy 627 AGTGGAGCTGGCTGGCGGCACTTTCATCATCCATCAGCGCGCAGGACTTTGGCAA 686
Db 918 AGCCGAGCTGCGGCTGGCAGACAACTTTCATCGCTCGGTGGCGCGCGCGACTTGGCAA 977

Qy 687 CATCAGCGGCTGTGTGGACCTTGACCTTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
Db 978 CATCAGAGGCTGTGTGATCTTGAGCCCTGTGCGGGAACACCATCGCCACGTGGTGTCCGG 1037

Qy 747 TTCCTTTTGTGACCTCGAGAGCCCTCCGCTCCCTCATCTTTCACAGCAATCGGCTGCCAAG 806

```

Db	2107	ATCCAGTACAAACAGCTCGGCTGATGACATCTCGTCTACAGGATGATCCCGCGGAGAGC	2166
Qy	1885	AAGGCCCTTCGTGGTCAACAACTGGTGTGAGGACTGGCTACGACTTGTGTGCTGGCC	1944
Db	2167	CGCTCGTTCCTGCTGACGGACCTGGGCTCAGGCGCGGACCTACGATCTGTGCGTGTCTCGCC	2226
Qy	1945	ATGTGGATACACAGCCAGCACATCAGCGCCACCAACATCGTGGCTGGCGCCAGTTTC	2004
Db	2227	GTGTATAGAGACAGCGCCACGGGGCTCACGGCCACGGGGCTGTGGCTGGCGCCGCTTC	2286
Qy	2005	TTCAACAAAGCTGACTACCTACCCGCACTGCGATCTCCATGCAACGACAGATTCTGGSCGCGACC	2064
Db	2287	TCACCGAAGACTGCGCTGCGGCATGCGGGCGCCGACGCTCCCTTCCTGGCGCGCAAG	2346
Qy	2065	ATGATCCTGCTATFCGGGGGATCATCGTGGCCACGCTGTGTTTCATCTCATCTTC	2124
Db	2347	ATGATCATCGCTGGCGGGCTCATCGTAGCTCTGGTACTGGTCTTCATCTTCGTGCTG	2406
Qy	2125	ATGGTGGCTACAAAGTCTGCAACACGAGGCCCCAG-----CAAGATGGCAGCGGCC	2178
Db	2407	CTAATGCGCTACAAAGGTGCACGGCGGCACGCCCCCGGCAAGGCAAGATTC	2466
Qy	2179	GTGAGCAATGTGTACTCGACACAAAGCGGCCCGGACACCGCTTCCAAAGCAGCACCA	2238
Db	2467	GTTAGCAGCTTTGCTCCAGACCAACGGCGGCCCTTGGGCCCGCCACGCGCCGCGG	2526
Qy	2239	GCCGGGGCCCCCGCCAGG	2257
Db	2527	CCCGCCCCGAGCGCCGCGG	2545
RESULT 15			
US-10-245-730-99			
; Sequence 99, Application US/10245730			
; Publication No. US20030073194A1			
; GENERAL INFORMATION:			
; APPLICANT: Baker, Kevin			
; APPLICANT: Baton, Dan			
; APPLICANT: Filvaroff, Ellen			
; APPLICANT: Goddard, Audrey			
; APPLICANT: Grimaldi, J. Christopher			
; APPLICANT: Gurney, Austin			
; APPLICANT: Smith, Victoria			
; APPLICANT: Stephan, Jean-Phillippe			
; APPLICANT: Watanabe, Colin			
; APPLICANT: Wood, William			
; APPLICANT: Zhang, Zemin			
; APPLICANT: Fong, Sherman			
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
; FILE REFERENCE: P36301C85			
; CURRENT APPLICATION NUMBER: US/10/245,730			
; CURRENT FILING DATE: 2002-09-16			
; PRIOR APPLICATION NUMBER: 10/197942			
; PRIOR FILING DATE: 2002-07-18			
; PRIOR APPLICATION NUMBER: 60/059114			
; PRIOR FILING DATE: 1997-09-17			
; PRIOR APPLICATION NUMBER: 60/063046			
; PRIOR FILING DATE: 1997-10-24			
; PRIOR APPLICATION NUMBER: 60/065027			
; PRIOR FILING DATE: 1997-11-10			
; PRIOR APPLICATION NUMBER: 60/079689			
; PRIOR FILING DATE: 1998-03-27			
; PRIOR APPLICATION NUMBER: 60/086478			
; PRIOR FILING DATE: 1998-05-22			
; PRIOR APPLICATION NUMBER: 60/087607			
; PRIOR FILING DATE: 1998-06-02			
; PRIOR APPLICATION NUMBER: 60/089801			
; PRIOR FILING DATE: 1998-06-18			
; PRIOR APPLICATION NUMBER: 60/090557			
; PRIOR FILING DATE: 1998-06-24			
; PRIOR APPLICATION NUMBER: 60/090689			
; PRIOR FILING DATE: 1998-06-25			

```

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
; LENGTH: 2855
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-245-730-99

Query Match          19.8%; Score 621.8; DB 9; Length 2855;
Best Local Similarity 61.3%; Pred. No. 1.3e-162; Indels 19; Gaps 4;
Matches 1078; Conservative 0; Mismatches 662;

Qy 507 GTTTCGGTGTGCGAGCTGCCCAAGTACTGTGCTTGCAGAAATCTGTCTGAGTCACT 566
    |||||
Db 798 GTCAGCACACCCAGCCATGTCCCGCGGTGCCGTGCCAGACACAGTGCCTGCCCT 857

Qy 567 GGGGACCTGTGCCCCCTCCAAAGGCTCTCTTTGTATCCCCCTGATATTACCGCGGAC 626
    |||||
Db 858 AAGCGTGTGTGCCAGGGGCGAGGCTCTCTGTTCGTGTCACCTCGCTGAGCCCGCGGC 917

Qy 627 AGTGGAGCTCGCCTGGGCGGCAACTTCATCATCCACATCAGCGCGCAGGACTTTGCCAA 686
    |||||
Db 918 AGCCGAGCTCGCGCTGGCAGACAACTTTCATCCCTTCGTGCGCCGCGGACCTGGCAA 977

Qy 687 CATGACGGGCTGGTGGACCTGACCCCTGTCCAGGAACCAATCAGCAGATCTCAGCCCTT 746
    |||||
Db 978 CATGACAGCGCTGCTGCATCTGAGCCTGTGCGGAACACCAATCCGCCACGTGGTGC CG 1037

Qy 747 TTCCTTTCTGGACTCGAGAGCTCGCTCCCTGCTGATCTTGACGCAATCGCTGCCAAG 806
    |||||
Db 1038 CGCCTTCGCCGACTCGCGGCCCTGCGTCCCTTGACCTGGATGCGACCGCTGACCTC 1097

Qy 807 CTTTGGGGAGGACACCCCTCCGGGCCCTGTGTCAACTCGACGACCTTATCGTGAAACAA 866
    |||||
Db 1098 ACTGGGAGGGCCAGCTGCGCGCCTGTGTCAACTTGCGCCACCTCATCTCAGCAACAA 1157

Qy 867 CCAGCTGGGCGGATCGCAGATGAGCTTTTGAGGACTTCTGCTGACATCGAGGATCT 926
    |||||
Db 1158 CCAGCTGGCAGCGTGGCGCGCGGCCCTGGATGATTTGTCGAGACACTGGAGGACT 1217

Qy 927 GGACCTCTCTACAAACACTCCATGCCCTGCCGTGGGACTCCGTGCGACGATGTCAA 986
    |||||
Db 1218 CGACCTCTCTACAAACACTCGAGCAGCTGCCCTGGAGGCCCTGGCGCGCTGGGCAA 1277

Qy 987 CTTCCACGAGTGAAGCTGGACACAACTGCTGTGATACATCGCCGAGGCACTTTGCG 1046
    |||||
Db 1278 CGTCAACACGTTGGGCCCTCGACACAACTGCTGGCTTCTGTGCC---GGCGCTTTTTC 1334

Qy 1047 AGACTGCAGAACTGCGCGCGCTGGATCTCACTCCAACTGGCTCGAGAAGCTGCCGCC 1106
    |||||
Db 1335 CGCGCTGCACAGCTGGCGCGCGCTGGACATGACCTCAACCGCTGACACAACTCCACC 1394

Qy 1107 TGATCCCATCTTTGCGCGCTCCGAGGCTTCGGCTTTGACAGCACACCTTTGCCGCCAC 1166
    |||||
Db 1395 CGACCCACTCTTCTCCGCGCTGCCCTGCTCGCGAGGCCCGGGGCTCGCGCGCTCTGC 1454

Qy 1167 CTTGTCTCTTAGTTTGGGGTACCCACTTCACCTGCAATTTGTAGCTTCTCTGGGTGCG 1226
    |||||
Db 1455 CTTGTCTCTGGCTTTGGCGGAAACCCCTGCTGCTGACCTGGAGCTGGTGGCTGCG 1514

Qy 1227 GAGGCTCGAGGGGACGATGACCTGGAACCTGTGGCTCCCGAGGGGCGCTCAAGGCTCG 1286
    |||||
Db 1515 TCGCTTGGCGGGAGGAGACCTCGAGGCTTCGCGCTGCCACTGCTCTGGGCGGCGCG 1574

Qy 1287 CTACTTCTGGCATGTGCTGAGGAGGAGTTTGTGTGGAGCGCGCTCTCATACCCAGCA 1346
    |||||
Db 1575 CTACTTCTGGGCGGTGGGCGAGGAGGTTTGTGTGGAGCGCGCGCTGTGTGACTCACCG 1634

Qy 1347 CACACAAAGTTGCTGGTTCTGTGGGGCGGAGCGCGCCACACTCAAGTGCAAAGCCATGG 1406
    |||||
Db 1635 CTCACACCTCTGGCTGTGCCCGCAGGTGCGCGGCTGCCCTGCGCTGCCGGGCGAGTGG 1694

Qy 1407 GGACCCGAGCCCTTTATCCACTGGGTAGCCCGCCGATGACCGCTGGTGGGAACCTCCTC 1466
    |||||

```

Search completed: June 8, 2003, 21:27:39  
Job time : 453 secs



---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 15:32:52 ; Search time 3910 Seconds  
(without alignments)  
13022.662 Million cell updates/sec

Title: US-09-831-846-1

Perfect score: 3144

Sequence: 1 gcctggctccctctcgtga.....gcgcttggtgattgcttatg 3144

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2170.6	69.0	3204	11	AK017594	AK017594 Mus muscu
2	609.4	19.4	611	14	BM787163	BM787163 K-EST0066
3	606.8	19.3	621	12	BF312768	BF312768 601896906
4	582	18.5	582	14	BQ126354	BQ126354 illeell.y
5	540.2	17.2	578	14	BQ126066	BQ126066 illeell.x
6	525.6	16.7	652	14	BM948451	BM948451 UI-N-EGOP

7	514.2	16.4	655	10	BB642310	BB642310
8	492.6	15.7	1853	11	BC007718	BC007718 Homo sapi
9	441.8	14.1	457	10	AW163387	AW163387 au94f03.y
10	422.8	13.4	435	10	AW161566	AW161566 au7le01.y
11	377.2	12.0	685	10	BB278209	BB278209
12	346.6	11.0	357	13	BM193100	BM193100 TCBAP1D95
13	327	10.4	635	12	EG380772	EG380772 UI-R-C10-
14	313	10.0	397	10	AW161909	AW161909 au7le01.x
15	310.6	9.6	708	13	BM046297	BM046297 603626095
16	300.4	9.6	714	14	BM950379	BM950379 UI-N-EHOP
17	293.8	9.3	674	14	BM722122	BM722122 UI-E-E00-
18	285.4	9.1	338	10	AW415047	AW415047 49056 MAR
19	279	8.9	563	10	AW274912	AW274912 xm62f02.x
20	276.6	8.8	712	14	BM944080	BM944080 UI-N-EHOP
21	273.4	8.7	822	14	BO571534	BO571534 UI-N-FCO-
22	271.4	8.6	703	14	BM784343	BM784343 K-EST0062
23	267.4	8.5	472	12	BF563759	BF563759 UI-R-BT1-
24	266.2	8.5	512	9	AI383513	AI383513 td03d01.x
25	266.2	8.5	513	9	AI418838	AI418838 lf37604.x
26	264.8	8.4	526	10	AW003909	AW003909 wq84a05.x
27	264.4	8.4	720	14	BM948138	BM948138 UI-N-EGOP
28	263.2	8.4	635	14	BM742719	BM742719 K-EST0015
29	263	8.4	513	9	AI760811	AI760811 w167g10.x
30	260.6	8.3	634	9	AI830569	AI830569 wj51c10.x
31	259.8	8.3	618	14	BM744495	BM744495 K-EST0018
32	259.8	8.3	620	14	BM742162	BM742162 K-EST0014
33	259.8	8.3	620	14	BM742315	BM742315 K-EST0015
34	259	8.2	1091	13	BM552976	BM552976 AGENCOURT
35	258.2	8.2	609	14	BM742747	BM742747 K-EST0015
36	257.6	8.2	609	14	BM742740	BM742740 K-EST0015
37	257.4	8.2	751	12	BE879258	BE879258 601484911
38	252.4	8.0	602	14	BM741814	BM741814 K-EST0014
39	252	8.0	618	14	BM744466	BM744466 K-EST0018
40	249.6	7.9	541	12	BF411974	BF411974 UI-R-BT1-
41	249.6	7.9	589	14	BM741945	BM741945 K-EST0014
42	245.4	7.8	600	14	BM949682	BM949682 UI-N-EGOP
43	243.8	7.8	581	14	BM783370	BM783370 K-EST0061
44	243.6	7.7	526	9	AA527134	AA527134 ni07a02.s
45	243	7.7	637	14	BM784328	BM784328 K-EST0062

ALIGNMENTS

RESULT 1	AK017594	3204 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK017594	Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730420005;homolog to KIAA1246 PROTEIN (FRAGMENT), full insert sequence.			
DEFINITION	AK017594	HTC; CAP trapper.			
ACCESSION	AK017594.1	GI:12856913			
VERSION					
KEYWORDS		Mus musculus (strain:CS7BL/6J) 8 days embryo cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library			
SOURCE		clone:5730420005.			
ORGANISM		Mus musculus			
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE					
AUTHORS		1			
TITLE		Carninci, P. and Hayashizaki, Y.			
JOURNAL		High-efficiency full-length cDNA cloning			
MEDLINE		Meth. Enzymol. 303, 19-44 (1999)			
PUBMED		99279253			
REFERENCE		2			
AUTHORS					
TITLE		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
JOURNAL		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
MEDLINE		Genome Res. 10 (10), 1617-1630 (2000)			
PUBMED		20499374			



QY	1571	CGGCCATGTTGGAGGTTCTCATCTGTTCAGCTGTCACACCTTACAGCAACAGCACAGCCGCGCA	1630
DB	1571	CGGCCACTGTAGAGGTTCTCCATTTGTGCAGCTCCACACCTTAGTAAACAGCACTAGCCGGA	1630
QY	1631	CTGCACCCCCCAAGTCCCGGCTCTCAGACATCACTGCTCCAGCAACAGCACCGGGGAG	1690
DB	1631	TGGCACCCCCCAAGTCTCGGCTTTTCGACATCACAGGTTCCAGCAACAGCACCGCGGGAG	1690
QY	1691	GTGAGGACAGTGGGGGCGGAGAGCTCCCAAAAGCCCCCGGAAACGGGCTGTGTTGTGT	1750
DB	1691	GGGAGGTAGTGGGCTGGGAGGCTCCCAAAAGCACCCACAGAGGGCTGTGTTGTGT	1750
QY	1751	CTGAAGTGACACACCTCGGCCCTGTGTAAGTGTGTCTAGCAAGTCAGCACCCCGGG	1810
DB	1751	CCGATGTTAGCACACCATCTGCCCTGTCAAGTGTCTGTCTGCAAGTCAGCACCCAGGG	1810
QY	1811	TGAAGATGATACAGTCGAGTACAACCTGTTCTGATGATGAGGTGCTGATCTTACAGTAGA	1870
DB	1811	TCCCAGGCTCCAAACAAGGCTTCGTGTGTAACAACCTGTGTCTCAGGACCTGGCTACGACT	1930
QY	1871	TCCCAGGCTCTACAAGGCTTCGTGTGTAACAACCTGTGTCTCAGGACCTGGCTACGACT	1930
QY	1931	TGTTGTGTGGGCATGTGGATGACACAGCCAGCACCTCAGGGCCACCAACATCTGTGG	1990
DB	1931	TGTCGGTACTGGCTATGTGGGACGACACAGCCAGCCAGCGCTCACAGCCACCAACATTTGG	1990
QY	1991	GCTGCGGCCAGTTCTTACCAGGCTGACTACCCGAGTGCAGTCCAGTCCATGSCACAGCCAGA	2050
DB	1991	GCTGTGCCCCAGTTCTTACCAGGCGGACTACCCCAATGCCAGTCTATGCACAGTCCAGA	2050
QY	2051	TTCTGGGGCGGACCATCATCTGCTATCGGGGCATCATCTGTGGCCACACGCTGCTGTCT	2110
DB	2051	TAAAGGGGGGACCATGATTTCTGGTCTATCGAGGTATCATTTGTGGCCACACTGCTGTCT	2110
QY	2111	TCATCGTCACTCTCATGGTTCGCTACAAGGTCTCAAGTCTGCAACACGAGGGCCCCCAGCAAGATGG	2170
DB	2111	TCATTGTAATCCTCATGGTTCGCTACAAGGTCTGCAACACGACACCCACAGGAAAAATGG	2170
QY	2171	CAGCGGCC----GTGAGCAATGTGTACTTCGCAGACCAACGGGGCCAGCCACCGGCTCAA	2227
DB	2171	CAGCGGCCACTGTGACGAATGTGTACTTCGCAGACCAATGGGTCTCAGCCCTCCACCTCTGG	2230
QY	2228	GCAGCGCACACGCGGGGCCCCGCGCGAGGGCCGCGGAAAGTGTGGTCGCGCAACGAGC	2287
DB	2231	GTGGGATACCGTTGGCCAGCTTCTCTCAGGCAACACCCAAAGTTGTGGTGGGAATGAGT	2290
QY	2288	TCCTGGACTTCACGCGCAGCTGCGCCGCGCCAGTGACTCTCTCTCCAGTCTCCCTGG	2347
DB	2291	TGATGGACTTCAGTACCAGCTGGCCAGGGCCCTGTGACTCTCTTCTTCAGGCTCACTGG	2350
QY	2348	GCAGTGGGAGGCTGCGGGGCTGGGACGGGCCCTCTGGAGGATCCACCTCTCGCCCCCGC	2407
DB	2351	GCACGGGGGAAGCGCGGGCTGGCGGGGTCCCTGTGAGGCTCCCAACCCCAAGCTCCTC	2410
QY	2408	GCCCAAGCCAGCTTGACCGCTGATGGGGCTTTCGCTCCCTGGACCTCAAGAGTC	2467
DB	2411	GTCCCAAGCCAGCTTGACCGCTGATGGAGCCTTTGGCTCCCTGGACCTCAAGAGTC	2470
QY	2468	AGAGAAAGGAGGAGCTGCTGACTCCAGGACTCCAGCGGGAGAGGGGCTGGGACGTCGG	2527
DB	2471	AAAGGAAGAAGAACTTCTAGACTCCAGGACTCCAGCAGCAGAGGGGAGGACATCAT	2530
QY	2528	CCGCGGGCCACACTCGGAACGAGAGCCACTGTGTGGGGCCCCCTGTGGGCCCGGGGCCAGGA	2587
DB	2531	CCAGAGGCCACCACTCAGACCGAGAGCCAATGCTGGGGGCCCCCGGGCCACCCGCTGCCAGGA	2590
QY	2588	GCTGTCTCCCTTGGCGTTGGAGGGCAAGGCCAAACGACAGGCACTCTTTCGACATGGGG	2647
DB	2591	GTCTTCTCCCTTGGCCCTTGGAGGGCAAGGCCAAACGAGGCCACTCTTTTGAATGGGGG	2650
QY	2648	ACTTTGTGCTGCGGGCGGGAGGGGTCTGTGCGGGCGGCTACAGTCTCTCTCGGAAGG	2707



D**b** 361

Query Match 18.5%; Score 582; DB 14; Length 582;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-114; Indels 0; Gaps 0;  
 Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1408 GACCCAGCCCTTATCCATCTGGGTAGCCCGGATGACCGCGTGTAGGAACTCTCA 1467  
 Db 1 GACCCAGCCCTTATCCATCTGGGTAGCCCGGATGACCGCGTGTAGGAACTCTCA 60  
 1468 AGACCGCTGTATGACAAATGGACCTTGGACATCTTCATCACCATCTCAGACAGT 1527  
 Db 61 AGACCGCTGTATGACAAATGGACCTTGGACATCTTCATCACCATCTCAGACAGT 120  
 1528 GTGCGCTTCACTGCAATCTCTGCAATGCTGCGGAGAGGCCACCGCATGATGAGGTC 1587  
 Db 121 GTGCGCTTCACTGCAATCTCTGCAATGCTGCGGAGAGGCCACCGCATGATGAGGTC 180  
 1588 TCCATCGTCAGTGCACACCTCAGCAACAGCAGCCGCTGACCCGCCCAAGTCC 1647  
 Db 181 TCCATCGTCAGTGCACACCTCAGCAACAGCAGCCGCTGACCCGCCCAAGTCC 240  
 1648 CGCTCTCAGACATCACTGCTCCAGCAAGACAGCCGCGGAGTGGAGCGAGTGGGGC 1707  
 Db 241 CGCTCTCAGACATCACTGCTCCAGCAAGACAGCCGCGGAGTGGAGCGAGTGGGGC 300  
 1708 GGAGAGCTCCCAAGAGCCCGGAGAGCGGCTGTGCTGTGCTGAGTGACCACTCC 1767  
 Db 301 GGAGAGCTCCCAAGAGCCCGGAGAGCGGCTGTGCTGTGCTGAGTGACCACTCC 360  
 1768 TCGCCCTGCTCAAGTGTCTGTCAGCAAGTCAAGCAGCCCGGCTGAAGTGTACCACTG 1827  
 Db 361 TCGCCCTGCTCAAGTGTCTGTCAGCAAGTCAAGCAGCCCGGCTGAAGTGTACCACTG 420  
 1828 CAGTACAACTGCTCTGACGATGAGTACTGATTTACAGATGATCCAGCTCCACAAG 1887  
 Db 421 CAGTACAACTGCTCTGACGATGAGTACTGATTTACAGATGATCCAGCTCCACAAG 480  
 1888 GCCTTGTGTCACAACTGCTGTCAGGAGTGGCTGACACTTGTGTGCTGGCCATG 1947  
 Db 481 GCCTTGTGTCACAACTGCTGTCAGGAGTGGCTGACACTTGTGTGCTGGCCATG 540  
 1948 TGGGATGACAGCCAGCAGCACTCACGGCCACCAACATCGTG 1989  
 Db 541 TGGGATGACAGCCAGCAGCACTCACGGCCACCAACATCGTG 582  
 RESULT 5  
 B0126066/c  
 LOCUS  
 DEFINITION  
 B0126066 578 bp mRNA linear EST 19-APR-2002  
 i116ell.x1 Melton Amplified Mouse E16 5 Pancreas 3 M16S1 A Mus  
 musculus cDNA clone IMAGE:5941749 3' similar to TR:Q9ULH4 Q9ULH4  
 KIAA1246 PROTEIN ; mRNA sequence.  
 B0126066  
 VERSION  
 B0126066.1 GI:20199977  
 KEYWORDS  
 SOURCE  
 house mouse.  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.  
 1 (bases 1 to 578)  
 Melton.D., Brown.J., Kenty.G., Permutt.A., Lee.C., Kaestner.K.,  
 Lemishka.I., Searce.M., Brestelli.J., Gradwohl.G., Clifton.S.,  
 Hillier.L., Marra.M., Pape.D., Wylie.T., Martin.J., Blistain.A.,  
 Schmitt.A., Theising.B., Ritter.E., Ronko.I., Bennett.J., Cardenas  
 ,M., Gibbons.M., McGann.R., Cole.R., Tsagareishvili.R., Williams.T.,  
 Jackson.Y. and Bowers.Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812

Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. Douglas Melton DNA sequencing by:  
 Washington University Genome Sequencing Center This clone is  
 available royalty-free through LNL; please contact the IMAGE  
 consortium (info@image.llnl.gov) for further information  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 388.  
 Location/Qualifiers  
 1..578.  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5941749"  
 A.  
 /sex="Both"  
 /tissue\_type="pancreas"  
 /dev\_stage="Embryonic day E16.5"  
 /lab\_host="DH10B"  
 /notes="Vector: pSPORT1; Site\_1: Not 1; Site\_2: Sal 1;  
 Library constructed using SuperScript plasmid library kit  
 (Life Technologies). cDNA made by oligo-dT priming.  
 Size-selected by column fractionation; average insert size  
 0.97 kb. Amplified once on solid support. cDNA Library  
 Preparation: Guolin Chen."  
 BASE COUNT 121 a 182 c 161 g 114 t  
 ORIGIN  
 Query Match 17.2%; Score 540.2; DB 14; Length 578;  
 Best Local Similarity 98.4%; Pred. No. 6.9e-105;  
 Matches 556; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
 QY 2575 GCCCGGCCAGGAGCGCTGCTCCCTTCCGTTGGAGGGCAAGCCCAACGACCACTCC 2634  
 Db 578 GCCCGGCCAGGAGCGCTGCTCCCTTCCGTTGGAGGGCAAGCCCAACGACCACTCC 519  
 QY 2635 TTCGACAT GGGGACATTTGCTGCTGCGGGCGGGAGGGTCTGTCGGGCGGCTACAG 2693  
 Db 518 TTCGACATGGGGGACATTTGCTGCTGCGGGCGGGAGGGTCTGTCGGGCGGCTCCAG 459  
 QY 2694 TCCTCCTCGAAGGTCTCGAATCTGACGAAAGCGGAGCTCTCTGTCACGCGCATGCT 2753  
 Db 458 TCCTCCTCGAAGGTCTCGAATCTGACGAAAGCGGAGCTCTCTGTCACGCGCATGCT 399  
 QY 2754 CTTGCGCTTTGAGGAGATGACCTGTTGGGGGCCCGGGGACTTTTGGCAGCTCCGAATG 2813  
 Db 398 CTTGCGCTTTGAGGAGATGACCTGTTGGGGGCCCGGGGACTTTTGGCAGCTCCGAATG 339  
 QY 2814 GGTGATGAGAGCACGGTCTAGGTGGGGTGGGCAATGCTCCCTTTCTGTGGGAGGCTG 2873  
 Db 338 GGTGATGAGAGCACGGTCTAGGTGGGGTGGGCAATGCTCCCTTTCTGTGGGAGGCTG 279  
 QY 2874 GGAGAGGGGAAAGAAATCTCACTGCAAGTGTGTTGGAGTTTCCATGATGTTTACA 2933  
 Db 278 GGAGAGGGGAAAGAAATCTCACTGCAAGTGTGTTGGAGTTTCCATGATGTTTACA 219  
 QY 2934 TCCAGGACAGTTTCTCTCTCCCTGTCATGGCTCGTGTCCCTTACCCCGCAACACC 2993  
 Db 218 TCAAAAACAGTTTCTCTCTCCCTGTCATGGCTCGTGTCCCTTACCCCGCAACACC 159  
 QY 2994 CACATCACTCCCCACCAACCCCGCGGGTGTGCTCAGGAAATGTCGCTCGCTCAATG 3053  
 Db 158 CACATCACTCCCCACCAACCCCGCGGGTGTGCTCAGGAAATGTCGCTCGCTCAATG 99  
 QY 3054 CCGGACTGAGCCCTGAGTGTGTTGGAAGCGGAGACTCCCGCTTTCTTAATCACAATGTAG 3113  
 Db 98 CCGGACTGAGCCCTGAGTGTGTTGGAAGCGGAGACTCCCGCTTTCTTAATCACAATGTAG 39  
 QY 3114 CCTACAGCAAGCGGCTTTGGATTG 3138  
 Db 38 CCTACAGCAAGCGGCTTTGGATTG 14



```

RESULT 6
BM948451          652 bp      mRNA      linear      EST 14-MAR-2002
LOCUS             UI-M-EG0p-bvd-m-12-0-UI.r1 NIH BMAP_EG0p Mus musculus cDNA clone
DEFINITION        IMAGE:5690675 5', mRNA sequence.
ACCESSION         BM948451
VERSION           BM948451.1 GI:19432042
KEYWORDS          EST.
SOURCE            house mouse.
ORGANISM          Mus musculus
Eukaryota; Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 652)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyX-5.
Location/Qualifiers
1..652
/organism="Mus musculus"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5690675"
/clone.lib="NIH BMAP_EG0p"
/tissue_type="whole brain"
/dev_status="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pyX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is CAGCCACGAC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institute of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT       143 a 183 c 189 g 137 t
ORIGIN
Query Match      16.7%; Score 525.6; DB 14; Length 652;
Best Local Similarity 87.9%; Pred. No. 9.2e-102;
Matches 573; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 1221 GCTCGGAGGCTCGAGCGGACGATGACCTGGAACCTGTGGCTCCCGAGGGGGCTCAA 1280
DB 1 GCTCGGAGGCTGGAGGGATGATGACCTGAGACCTGTGGATCCCGGAAGTCTCAA 60
QY 1281 GGGTCGCTACTTCTGGCATGTCGCTGAGGAGAGTTGTGCGAGCGCCCTCTCATCAC 1340
DB 61 GGGTCGCTACTTCTGGCATGTCGCTGAGGAGAGTTGTGCGAGCGCCCTCTCATCAC 120
QY 1341 CCAGCACACACAACTTCTGGTTCTGGAGGGCCAGGGCCACACTCAAGTGAAGGC 1400
DB 121 CCAGCACACACAACTGCTGTTCTGGAGGGCCAGGGCCAGGCACCTTTAAGTGAAGGC 180
QY 1401 CATTGGGGACCCAGCCCTTATCCACTGGGTAGCCCGCCGATGACCCGCTGGTAGGAA 1460

```

```

Db 181 CATTGGGACCCAGCCCTTTGATCCACTGGTGTGCCCCCGATGACCCCTTGGTGGGAAA 240
QY 1461 CTCCTCAAGGACCGCTGTCTATGACAAATGGACACCTCGGACATCTTTCATCACCACATCTCA 1520
Db 241 CTCCTCCAGGACTGCTGTATATGACAAATGGACACCTCGGACATCTTAAATACCACTCTCA 300
QY 1521 GGACAGTGGTGCCTTACCTGCTGCTGCAATGCTCCGGAGAGGCCACGGCCATGGT 1580
Db 301 GGACAGTGGACCTTTTACCTGCTGCTGCAATGCTCCGGAGAGGGTACGGCCACTGT 360
QY 1581 GGAGTCTCCATCGTCCAGCTGCCACACTCAGCAACAGCAGCAGCCCGATGACACCC 1640
Db 361 AGAGTCTCCATGTCAGCTGCCACACTTAGTAACAGCAGTACGCCGATGGCAGCC 420
QY 1641 CAAGTCCCGCTCTCCAGACATCACTGCTCCAGCAAGACCCAGCGGGAGGTGGAGGAG 1700
Db 421 CAAGTCTCGCTTTCGGACATCACTGCTCCAGCAAGACCCAGCGGGAGGTGGAGGAG 480
QY 1701 TGGGGCGGAGAGCTTCCAAAGCCCGGAGCGGCTGTGCTGTGTCTGAAGTGCAC 1760
Db 481 TGGGGCTGGGAGCTTCCAAAGCAGCAGAGAGGGCTGTGCTGTGTGCTCGATGTTAC 540
QY 1761 CACCACTCGGCGCTGCTCAAGTGTCTGAGCAAGTCAAGCAGCCCGGCTGAAGATGA 1820
Db 541 CACCACATCTCGCTGCTCAAGTGTCTGAGCAAGTCAAGCAGCCCGGCTGAAGATGA 600
QY 1821 CCAGCTGCAGTACAACTGCTCTGAGGATGAGTACTGATTTACAGGATGATC 1872
Db 601 TCAGCTGCAGTACAACTGCTCTGAGGATGAGTACTGATTTACAGGATGATC 652

RESULT 7
BB642310          655 bp      mRNA      linear      EST 31-AUG-2001
LOCUS             BB642310 RIKEN full-length enriched, adult retina Mus musculus CDNA
DEFINITION        clone A930009011 5', mRNA sequence.
ACCESSION         BB642310
VERSION           BB642310.1 GI:15401853
KEYWORDS          EST.
SOURCE            Mus musculus
ORGANISM          house mouse.
REFERENCE          1 (bases 1 to 655)
AUTHORS           Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M.,
Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okada,Okazaki,Y.,
Okido,T., Saito,R., Sakai,C., Sakai,K., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M.,
Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M.
and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M.,
Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S.,
Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format

```





Tue Jun 10 08:57:38 2003

Email: est@watson.wustl.edu  
This clone is available royalty-free through LINL ; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Seq primer: -40RP from Gibco  
High quality sequence stop: 431.

FEATURES

1.1.435  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2781720"

BASE COUNT	83 a	131 c	140 g	103 f
100	100	100	100	100
101	101	101	101	101
102	102	102	102	102
103	103	103	103	103
104	104	104	104	104
105	105	105	105	105
106	106	106	106	106
107	107	107	107	107
108	108	108	108	108
109	109	109	109	109
110	110	110	110	110
111	111	111	111	111
112	112	112	112	112
113	113	113	113	113
114	114	114	114	114
115	115	115	115	115
116	116	116	116	116
117	117	117	117	117
118	118	118	118	118
119	119	119	119	119
120	120	120	120	120
121	121	121	121	121
122	122	122	122	122
123	123	123	123	123
124	124	124	124	124
125	125	125	125	125
126	126	126	126	126
127	127	127	127	127
128	128	128	128	128
129	129	129	129	129
130	130	130	130	130
131	131	131	131	131
132	132	132	132	132
133	133	133	133	133
134	134	134	134	134
135	135	135	135	135
136	136	136	136	136
137	137	137	137	137
138	138	138	138	138
139	139	139	139	139
140	140	140	140	140
141	141	141	141	141
142	142	142	142	142
143	143	143	143	143
144	144	144	144	144
145	145	145	145	145
146	146	146	146	146
147	147	147	147	147
148	148	148	148	148
149	149	149	149	149
150	150	150	150	150
151	151	151	151	151
152	152	152	152	152
153	153	153	153	153
154	154	154	154	154
155	155	155	155	155
156	156	156	156	156
157	157	157	157	157
158	158	158	158	158
159	159	159	159	159
160	160	160	160	160
161	161	161	161	161
162	162	162	162	162
163	163	163	163	163
164	164	164	164	164
165	165	165	165	165
166	166	166	166	166
167	167	167	167	167
168	168	168	168	168
169	169	169	169	169
170	170	170	170	170
171	171	171	171	171
172	172	172	172	172
173	173	173	173	173
174	174	174	174	174
175	175	175	175	175
176	176	176	176	176
177	177			

```

Query Match      14.1%; Score 441.8; DB 10; Length 457;
Best Local Similarity 99.3%; Pred. NO. 6.7e-84;
Mismatch 0; Mismatches 2; Indels 1; Gaps 1;
Watch 454; Conservative

```

66 ACTTGAAGACACCCACATTCCAAGATGCCCGAGGTTCTGGGAAATCCATGGGATTTTC 60

db  
1 ACTTGAAGACACCCCATTC

126 GATCCGGAATACTACCGCATCCTAGGGAGGATATATATTATTTCTT

CTCCGCAAAATCCTACCGGCATCCTCTAGGAGGATTATTATTATTTCTTA 120

...GCTTCTTTCCCTTCTTCTTAACGCCATGG 245

[illegible]

D<sub>b</sub> 121 ATCTGGAAGAGAAGACAAAGTGGAGCTTACCGGCTT

246 ATATAACTGAAT-AAGCGGCTCAGGGCTTTCCCGCGTGGACGTCGAGCCGCGGCTT

181 ATATAACTGAATAAAGCCGCTCAGGGCTTCCCCCGGTGACGTCCGAGGCCACCACTCG 240

364

300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098

Db  
241 CCIGCAATCGCCGAGGCCGGGCTTTC

Oy  
365 GCGTGGCCGAGCCGGGGAGCCGGGGGCCCGGACCGGC

nb. 301 GCGTGGCCGAGCCGGGGAGCCCGGGCGCCCGGGAGCCGGCCCTCGGTGCGCCAGCCG

5' CCGCTACATGCTGCCTCGCCCGGCTGAGTGACCAAGACCATGGAGACCTGCTTGGTG 484

[illegible][illegible]

QY  
485 GCCCTGCTAGCGTTGGCAAGCCGTTCGCCCTGCCTCCTT

nb  
421 GCCTGCTAGCGTTTGGCATGGCGTTTGCCCGTGGT'CGA 457.

## RESULT 10

AW161566	AW161566	435 bp	mRNA	linear	EST: U9-NCV-139
LOCUS	AW161566				

**DEFINITION** au71e01.y1 Schneider Lectar strain  
 IMAGE:2781720 5' mRNA sequence.

ACCESSION AW161566

```

VERSION
*****
AWI61300.1 01.03.000000
*****
EST
*****

```

SOURCE human.

Eukaryota; Metazoa; Chordata; Vertebrata; Eucercoscomia

Order	Family	Genus	Species	Number	Sex	Age	Weight	Length	Wing	Tail	Head	Bill	Tarsus	Middle toe	Claw
				1	♂	Ad	1.435	190	110	75	25	15	25	15	10

**AUTHORS**  
 Hillier, L., Allen, M., Bowles, L., Dubuque, J., Gosselin, J., Marra, M., Lennon, G., Iacov M.

J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Ingersing, D.,

White, Y., Wylie, I., MacCabe, J., Wray, N., & Wray, N. (2014). The WACHU-NCT human EST Project

JOURNAL Unpublished (1997)

COMMENT  
Other Eds: as above  
Contact: Wilson RK

Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

0  
CCT  
CCT  
HTC :XPB



1

RESULT 11  
BB278209  
LOCUS  
DEFINITION

	BB278209	685 bp	mRNA	linear	Mus musculus	CDNA
	BB278209	RIKEN full-length enriched,	adult retina			

	BB278209	685 bp mRNA	RIKEN full-length enriched, adult retina	Mus musculus	CDNA
	BB278209				



Tue Jun 10 08:57:38 2003

Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP9544, mRNA sequence.

ACCESSION BM193100  
VERSION BM193100.1 GI:17651296  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 357)

## AUTHORS

Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,  
Gunaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.  
Pediatric Leukemia cDNA Sequencing Project (2001)

## TITLE

Unpublished (2001)

## JOURNAL

Contact: Dr. Judith F. Margolin

## COMMENT

Texas Children's Cancer Center and Human Genome Sequencing Center  
at Baylor College of Medicine  
1102 Bates, MC3-3320 Houston, TX 77030, USA  
Tel: 832-824-4536  
Fax: 832-825-4038  
Email: clones@ccc.org

Seq primer: M13 primer

## FEATURES

Location/Qualifiers

1..357

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="TCBAP9544"

/clone\_lib="pediatric pre-B cell acute lymphoblastic

leukemia Baylor-HGSC project-TCBA"

/sex="male"

/tissue\_type="leukopheresis"

/cell\_type="pre-B cell"

/dev\_stage="pediatric 2 years"

/lab\_host="DH10B"

/note="vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;

First strand cDNA was primed with an anchored

3'-oligo(dT) primer [5'-GGAGACTCGAGCGCGCAGGAG(T)VN

3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand

was primed with a BamHI-dC primer

[5'-AGAGAGTCGATCGCGCGCAATAAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI

and directionally cloned into the BamHI and SalI sites of

lambda pSB vector. Library was constructed by Wei Yu at RIKEN

normalization. Library was constructed by Wei Yu at RIKEN

of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,

Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,

Schneider C, Hayashizaki Y, High efficiency selection of

full-length cDNA by improved biotinylated cap trapper..

DNA Res 4: 1, 61-6, Feb 28, 1997)"

71 a 111 c 103 g 72 t

## BASE COUNT

## ORIGIN

Query Match 11.0%; Score 346.6; DB 13; Length 357;

Best Local Similarity 98.9%; Pred. No. 1.3e-63;

Matches 349; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

## QY

1240 GACGATGACTGGAACCTGTGGCTCCCGAGGGCCCTCAAGGTCGCTACTTCTGGCAT 1299

## DB

1 GACGATGACTGGAACCTGTGGCTCCCGAGGGCCCTCAAGGTCGCTACTTCTGGCAT 60

## QY

1300 GTGCGTGAGGAGGAGTTGTGTGCGAGCGCGCTCTCATCCACCGACACACAAAGTTG 1359

## DB

61 GTGCGTGAGGAGGAGTTGTGTGCGAGCGCGCTCTCATCCACCGACACACAAAGTTG 120

## QY

1360 CTGTTCTGAGGGCGAGCGGCCACACTCAAGTGCAAGGCAATGGGACCCAGCCCC 1419

## DB

121 CTGTTCTGAGGGCGAGCGGCCACACTCAAGTGCAAGGCAATGGGACCCAGCCCC 180

## QY

1420 CTTATCCACTGGGTAGCCCCCGATGACCGCTCTGGTAGGGAATCTCTCAAGGACCGCTGTC 1479

## DB

181 CTTATCCACTGGGTAGCCCCCGATGACCGCTCTGGTAGGGAATCTCTCAAGGACCGCTGTC 240

## QY

1480 TATGACAATGGCCCTTGGACATCTTTCATCACCACATCTCAGGACAGTGGTGCCTTCACC 1539

Db 241 TATGACAATGGCCCTTGGACATCTTTCATCACCACATCTCAGGACAGTGGTGCCTTCACC 300

QY 1540 TGCATTGCTGCCAATGCTCCCGAGAGCCCGCATGGTGGAGGTTCTCCAT 1592

Db 301 TGCATTGCTGCCAATGCTCCCGAGAGCCCGCATGGTGGAGGTTCTCAT 353

## RESULT 13

## BG380772/c

## LOCUS

## DEFINITION

UI-R-CTO-bub-c-04-0-UI.s1 UI-R-CTO Rattus norvegicus cDNA clone

UI-R-CTO-bub-c-04-0-UI 3', mRNA sequence.

## ACCESSION

BG380772

## VERSION

BG380772.1

## KEYWORDS

EST.

## SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 635)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the

normalized rat brain pool library cDNA Library preparation: M.B.

Soares Lab Clone distribution: clones will be available through

Research Genetics (www.resgen.com)

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1..635

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-CTO-bub-c-04-0-UI"

/dev\_stage="ADULT"

/lab\_host="DH10B (Life Technologies)"

/note="Vector: pT73D-pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CTO

library is a normalized library constructed from the

following rat brain tissues: embryonic day 17, embryonic

day 19, embryonic day 21, adult day 1, adult day 12, adult

day 75, adult day 200. For a detailed description of the

library from which this clone was derived, please visit

our web site at ratseq.eng.uiowa.edu. The subtraction has

been previously described in (Bonaldo, Lennon and Soares,

Genome Research 6:791-806, 1996)

TAG LIB=UI-R-CTO

TAG TISSUE=rat brain pool

TAG SEQ=ACTTC

135 a 174 c 183 g 143 t

## BASE COUNT

## ORIGIN

Query Match 10.4%; Score 327; DB 12; Length 635;

Best Local Similarity 77.4%; Pred. No. 2.2e-59;

Matches 485; Conservative 0; Mismatches 110; Indels 32; Gaps 6;

```
QY 2533 GGCACCACTCGACGAGACCACTGCTGGGGCCCCCTGGCGCCGCGGCGCAGGACCTG 2592
Db 635 GGCACCACTCGACGAGACCACTGCTGGGGCCCCCAGCACCCGCGGCGCAGGACCTT 576
QY 2593 CTCCTCTGCGCTGGAGGGCAAGGCAACAGCAGCACTCTTCGACATGGGGACTTT 2652
Db 575 CTCCTCTGCGCTGGAGGGCAAGGCAACAGCAGCACTCTTCGACATGGGGACTTT 516
QY 2653 GCTGTGCGGGCGGCGGAGGGGTGCTGCGGGCGGCTACAGTCTCTCGGAAGGTCTGG 2712
Db 515 GCAGCTGCAGCTGC-----AGCTGCTCTGCTGGGTACAGTCCGCCCTAGGCGGGTCTCA 462
QY 2713 AACATCTGGACAGGCGAGCTCTCTGTCAAGGCATGCTCTGCCCTTTGAGGAGAT 2772
Db 461 AACATCTGGACAGGCGAGCTCTGTCTCAATGGCATGCTCTGCCCTTTGAGGAGAT 402
QY 2773 GACCTGGTGGGGCGGCGGAGCTTTTGGCAGCTCCGAATGGGTGATGGAGACGCGTC 2832
Db 401 GACCTGGTGGGAGCTCGGGGACATTTGGCAGCTCAGATGGGTATGGAGAGTACTGTG 342
QY 2833 TAGGTGGGGGTGGCATGCTCCCTTTCTGTGCGCAGGSGTGGGAGAGGGGAAAGATC- 2891
Db 341 TAGCCAGGGAAGGATACCGTCCCTCCCGCCCTTCTCAAAAGTGGGAGAGGAAGGATCA 282
QY 2892 -----TCACTGGCAAGTGTGTTGGAGTTTCATGTTGATGTTTACATCCAGGACAGT 2945
Db 281 GAGTCAGCACTGGCAAGTATTTGTGGAGTTTCCATGGTGTATGTTTACATCCAGGACAGT 222
QY 2946 TTCTGCTCCCTCTCAATGCCCTCGTGTG-----CCCCCTACCCCGCAACCCACATC 2999
Db 221 CTGTGCTCCCTCTCAACAGCTCATGTCCTTCTGCCCCAGCTACACCCCTTCCCTATGTC 162
QY 3000 ACCTCCCCA-----CCACCGCGCGGGGTGTGTCAGGGGAATGTGACTCGTCT 3048
Db 161 ACCTTCCAGGCGTGTCCCCACACCCAGCGGGTGTGCTCAGGGAATG-GGACTCGTCT 103
QY 3049 AATGCCGAGCTAGCCCTGAGTGTGTTGGAAGGCGAGACTCGGCTTTCTATCACAA 3108
Db 102 ATGTGTCACTGAGCGCTGAGTGTGTTGGAAGGCTTAACTGCTTTCTATTACAAA 43
QY 3109 TGTAGCTACAAGCAAGCGGCTTTTGA 3135
Db 42 TGTAG-CGACAAAGCAAGCGGCTTCGAA 17
```

```
RESULT 14
AW161909/c 397 bp mRNA linear EST 09-NOV-1999
LOCUS au7le01.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone
DEFINITION IMAGE:2781720 3', mRNA sequence.
ACCESSION AW161909
VERSION AW161909.1 GI:6300942
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 397)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Other ESTs: au7le01.y1
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
```

```
Seq primer: -40UP from Gibco
High quality sequence stop: 390.
Location/Qualifiers
1. 397
FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2781720"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue type="frontal lobe"
/lab_host="DH10B"
/dev stage="5 months post-conception"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site 1: SstI; Site 2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGAGAGAGAGAGCTCAAGGATCTTAATTAATTAATCCCCCCCCCCC-3'
and 3' adaptor sequence:
5'-GAGAGAGAGACTCGAGTCTTTTCTTTT-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (UNCIB-Area
Science Park, Trieste, Italy)."
BASE COUNT 84 a 113 c 91 g 100 t
ORIGIN
Query Match 10.0%; Score 313; DB 10; Length 397;
Best Local Similarity 90.6%; Pred. No. 2e-56; 34; Indels 0; Gaps 0;
Matches 328; Conservative 0; Mismatches 34;
QY 2781 GGGGGCCCCGGGGAGCTTTTGGCAGCTCCGAATGGGTGATGGAGACACGGTCTAGGTGG 2840
Db 395 GGGGGCCCCGGGGAGTCTTTTGGCAGCTCCCAATGGGTGATGGAAAGCCCGGTCTAGGTGG 336
QY 2841 GGTGGGATGCTCCCTTTCTGTGCGAGGTGGGAGAGGGAAGAATCTACTGGCA 2900
Db 335 GGTGGGCAAGCTCCCTTTCTGTGCGAGGTGGGAAAGGGAAGAAATTTTCACTGGCA 276
QY 2901 AGTGTGTTGGAGTTCATGTTTACATCCAGGACAGTTCCTCTCCCTGTCTCA 2960
Db 275 AGGTTTGGAGTTCATGTTTACACAGGAGCGGTTTGTCTCCCGNNA 216
QY 2961 ATGGCTCTGTGTCCTCCCTACCCGCAACACCCACATCACCTCCACACCGCCCG 3020
Db 215 AGGGCTGTGTCCTCCCGCNCNNAACCCACATCCCTCCCTCCCTCCCTCCCGCCG 156
QY 3021 GGTGTGCTCAGGAATGTGGACTCGCTCAATGCGGAGCTGAGCCCTGAGTGTGGAAA 3080
Db 155 GGTGTGCTCAGGAATGTGGATTCGCTCAATGCGGAGTTGAGCCCTGAGTGTGGAAA 96
QY 3081 GCGGAGACTCGGCTTCTTAATCAGAAATGTAGCTACAGCAAGCGGCTTTGGATTGCT 3140
Db 95 GGGGAGACTCCCTTTTAAATCAAAATGTAGCTTACAGCAAGGGGCTTTGGATTGTT 36
QY 3141 TA 3142
Db 35 TA 34
RESULT 15
BM046297
LOCUS 603626095F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5452630 5',
DEFINITION mRNA sequence.
ACCESSION BM046297
VERSION BM046297.1 GI:16775564
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 708)  
REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: DCTP/DTP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM1944 row: n column: 23  
High quality sequence stop: 704.  
Location/Qualifiers  
1. .708  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5452630"  
/clone\_lib="NIH\_MGC\_40"  
/tissue\_type="carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dt priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAGGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH MGC library."

BASE COUNT 101 a 272 c 209 g 126 t  
ORIGIN

Query Match 9.9%; Score 310.6; DB 13; Length 708;  
Best Local Similarity 65.8%; Pred. No. 7e-56;  
Matches 473; Conservative 0; Mismatches 234; Indels 12; Gaps 1;  
QY 651 CTTTCATCATCCACATCAGCGCCGAGGACTTCCACATGAGGGGCTGGTGGACCTGAC 710  
Db 2 CTTTCATCAGGGCCCTGGGGCCCTGACTTCGCAACATGACGGGACTGGTGGACCTGAC 61  
QY 711 CTTGTCAGGAAACACCATCAGCCATCCAGCCCTTTCTTCTGAGCTCGAGAGCCT 770  
Db 62 ACTGTCTCGAATGCCATCACCAGCATTTGGGGCCCGGCTTTGGGACCTCGAGAGCCT 121  
QY 771 CGCTCCCTGCAATCTTGACAGCAATCGGTGCGCAAGCCCTTGGGGAGGACACCCCTCGGGG 830  
Db 122 GCGTTCCTCCACTTGACGGCAACAGGCTGTGGAGCTGGGCAACCGGGAGCCTCGGGG 181  
QY 831 CTTGGTCAACCTGACGACCTTATCGTGAACACACAGCTGGGGGCGCATCGCAGATCA 890  
Db 182 CCCCCTCAATCTGACGACCTTATCTCAGCGGCAACAGCTGGGGCGCATCGGGCGGG 241  
QY 891 GGCTTTTGAGGACTTCTGTCACATTTGAGGATCTGAGACCTCTCTACAAACCTCCA 950  
Db 242 AGCCTTCGAGACTTCTAGAGAGCCTGGAGGACCTGAGACCTGCTCTACAAACCTCCG 301  
QY 951 TGGCCTGCGGTGGGATCCGTGGGAGCATGFTCAACCTCCACAGCTGAGCTGGAGCCA 1010  
Db 302 GCAGGTGCGGCTGGGGCGGATCGGGCCATGCTGCTGCTGCAACCTCAACCTGGACCA 361  
QY 1011 CAACCTGCTGGATCACATCGCGAGGGACCTTTGAGACCTCGAGAACTGGCGGCT 1070  
Db 362 TAACTTATGAGGACTGCCCCCAGGGCCCTTCGCCAGCTCGGTGAGCTCTCCGCT 421  
QY 1071 GGATCTACCTCCAAATCGGTGAGAAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCA 1130  
Db 422 GGACCTCACCTCAACCGGCTGSCCAGCTGGCTCCGAGCCGCTTTCTCTCGTGGGG 481  
QY 1131 GCGTTGGGCTTTGACAGGACACACCTTTGCCCCACCCCTTGTCTTTAGTTTGGGGTAA 1190

Db 482 TGATGCAGAGG-----CCTCTCCCGCCCCCTGTGTGTGAGCTTTAGCGGAA 529  
QY 1191 CCCACTTCACATGCAATTTGTGAGCTTCTCTGCTGCGGAGGCTCGAGCGGAGATGACCT 1250  
Db 530 CCCCCCTGCACTGCAAACTGTGAGCTGTGTGCTGCGGCGCTGGCGCGCGGAGCAGCT 589  
QY 1251 GGAAACCTGTGCTCTCCCGAGGGGCTCAAGGGTCTGCTACTTCTGGCATGTGCTGAGGA 1310  
Db 590 GGAAACCTGTGCTCTCCCGAGGGGCTCGCGGCGCTACTTCTGGGAGTGGCGGAGGG 649  
QY 1311 GGAGTTTGTGTCGAGCGGCTCTCATCACCAGCAGACACACAAGTTGCTGTTCTGG 1369  
Db 650 CGAGTTCTCTGTGAGCCTGCGCTCATTTGCCCGACACACGAGCGCCTCTGGGTGCTTG 708

Search completed: June 8, 2003, 19:02:56  
Job time : 3928 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 19:03:03 ; Search time 63 Seconds  
(without alignments)  
1668.805 Million cell updates/sec

Title: US-09-831-846-2  
Perfect score: 4094  
Sequence: 1 METLGLLAFQMAFVDA.....DLVGARTGSGSEWNWSTV 789

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4094	100.0	789	21 AAB09968	Human brain-specif
2	4087	99.8	789	22 AAU28092	Novel human secret
3	4087	99.8	789	22 AAM39059	Human polypeptide
4	4075	99.5	785	21 AAB12448	Human hh00149 prot
5	1846	45.1	766	22 AAG67505	Amino acid sequenc
6	1820.5	44.5	719	23 ABP43533	Human secreted pro
7	1703.5	41.6	771	22 AAG67512	Amino acid sequenc
8	1667.5	40.7	635	23 AAE23380	Human LP220 secret
9	1656	40.4	322	22 ABG13536	Novel human diagno
10	1586.5	38.8	636	22 AAU32870	Novel human secret

11	1573	38.4	628	22 AAB84469	Amino acid sequenc
12	1573	38.4	628	22 AAG67523	Amino acid sequenc
13	1566	38.3	628	22 AAG65805	Human leucine-rich
14	1563.5	38.2	627	23 ABG34079	Human pro peptide
15	1545.5	37.8	551	23 AAE17484	Human leucine-rich
16	1468.5	35.9	526	22 ABG04827	Novel human diagno
17	1319.5	32.2	533	22 AAG67511	Amino acid sequenc
18	1313.5	32.1	503	22 ABG05946	Novel human diagno
19	1241	30.3	768	22 ABG19413	Novel human diagno
20	1239	30.3	479	22 AAG67522	Amino acid sequenc
21	1144	27.9	468	22 AAB70072	Human secreted pro
22	1144	27.9	468	23 ABG65511	Human albumin fusi
23	721	17.6	160	22 AAM40845	Human polypeptide
24	651	15.9	136	22 AAU28280	Novel human secret
25	511	12.5	180	22 ABG03072	Human expressed po
26	511	12.5	180	22 AAU20510	Human secreted pro
27	452	11.0	713	23 AAU91335	Human novel secret
28	451	11.0	713	22 AAE13006	Human leucine-rich
29	443.5	10.8	656	23 AAU91341	Human novel secret
30	435.5	10.6	640	20 AAU08100	Human PRO331 prote
31	435.5	10.6	640	20 AAU85722	Novel protein (Cio
32	435.5	10.6	640	20 AAU13394	Amino acid sequenc
33	435.5	10.6	640	21 AAB24407	Human PRO331 prote
34	435.5	10.6	640	21 AAY70673	Human PRO331 prote
35	435.5	10.6	640	22 AAU12355	Human PRO331 poly
36	435.5	10.6	640	22 AAU00826	Human immune respo
37	435.5	10.6	640	22 AAB80262	Human PRO331 prote
38	435.5	10.6	640	22 AAB65292	Human PRO331 prote
39	435.5	10.6	640	22 AAB53089	Human angiogenesis
40	420.5	10.3	606	23 ABG60230	Human Slit-3-like
41	420.5	10.3	606	23 AAU85411	Human protein NOV9
42	419	10.2	606	23 ABG60229	Human Slit-3-like
43	419	10.2	606	23 AAU85410	Human protein NOV8
44	418.5	10.2	579	22 AAE06804	Mature human neuro
45	418.5	10.2	606	22 AAE06799	Human neuronal gui

ALIGNMENTS

RESULT 1  
AAB09968  
ID AAB09968 standard; Protein; 789 AA.

AC AAB09968;

DT 19-OCT-2000 (first entry)

DE Human brain-specific transmembrane glycoprotein.

XX Transmembrane glycoprotein; human; brain; cellular signal transducer;  
KW neuroactive; neuroprotective; cerebroprotective; drug development;  
KW treatment; nervous disease; diagnostic.

XX Homo sapiens.

OS WO2000031256-A1.

PD 02-JUN-2000.

PF 18-NOV-1999; 99WO-JP06449.

PR 20-NOV-1998; 98JP-0331727.

XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

PA (KAZU-) KAZUSA DNA RES INST.

PI Funahashi S, Miyata S, Nomura N, Nagase T, Ohara O;

XX WPI; 2000-411954/35.

DR N-PSDB; AAA40083.

XX Gene encoding a brain-specific transmembrane glycoprotein that has a

PT typical PDZ protein binding motif and functions as a cellular signal  
transducer, useful in developing drugs for treating nervous diseases -  
XX  
XX  
PS Claim 1a; Page 52-57; 63pp; Japanese.

XX This invention describes a novel gene encoding a human brain-specific  
transmembrane glycoprotein that has a typical PDZ protein binding motif  
and functions as a cellular signal transducer. The product of the  
invention has neuroactive, neuroprotective and cerebroprotective  
activity. The gene and encoded protein are useful in developing drugs  
for treating nervous diseases, and also for studying functions of the  
nervous system or onset mechanism of nerve-related diseases. They may  
also be used for the production of diagnostic reagents such as  
oligonucleotide probes and antibodies for detecting proteins comprising  
PDZ domains and the nucleic acids that encode them. They may also be  
used to produce and identify modulators of the proteins expression and  
activity such as antisense sequences and antibodies. This sequence  
represents the human brain-specific transmembrane glycoprotein described  
in the method of the invention.

XX Sequence 789 AA;  
Query Match 100.0%; Score 4094; DB 21; Length 789;  
Best Local Similarity 100.0%; Pred No. 1.8e-303;  
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 METLLGGLLAFGMAFAVVDACPKYCVQNLSESLGTLCPKSGLLFPVPPIDRRTVLERLG 60  
DB 1 METLLGGLLAFGMAFAVVDACPKYCVQNLSESLGTLCPKSGLLFPVPPIDRRTVLERLG 60  
QY 61 GNFIHISRODFANMTGLVDTLSRNTISHIQPSFLDLSLSLHLDNSNRLPSLGEDTL 120  
DB 61 GNFIHISRODFANMTGLVDTLSRNTISHIQPSFLDLSLSLHLDNSNRLPSLGEDTL 120  
QY 121 RGLVNLQHLI VNNNQLGGIADEAFEDLLTLEDLSDYNNLHGLPWSVRRMYNLHQLSL 180  
DB 121 RGLVNLQHLI VNNNQLGGIADEAFEDLLTLEDLSDYNNLHGLPWSVRRMYNLHQLSL 180  
QY 181 DHNLDHIAEGTFADLOKLARDLTSNRLOKLPDPPIFARSQASALTATFAPPLPSFG 240  
DB 181 DHNLDHIAEGTFADLOKLARDLTSNRLOKLPDPPIFARSQASALTATFAPPLPSFG 240  
QY 241 GNPLHCNCELLWLRLERDDLETCTGSPGLKGRFYHVRREEFVCEPPLITQHTHKLIV 300  
DB 241 GNPLHCNCELLWLRLERDDLETCTGSPGLKGRFYHVRREEFVCEPPLITQHTHKLIV 300  
QY 301 LEGQAATLKCAIGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360  
DB 301 LEGQAATLKCAIGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360  
QY 361 AANAAGEATAMVEISIVOLPHLSNSTSTAPPKSRLSDITGSSKTSRGGSGGGGEPKPS 420  
DB 361 AANAAGEATAMVEISIVOLPHLSNSTSTAPPKSRLSDITGSSKTSRGGSGGGGEPKPS 420  
QY 421 PPERAVLVSEVTTTSALVKWSVKSAPRVKMYQYQNCSDDEVLIYRMI PASNKAFVYNN 480  
DB 421 PPERAVLVSEVTTTSALVKWSVKSAPRVKMYQYQNCSDDEVLIYRMI PASNKAFVYNN 480  
QY 481 LVSGTG YDLCVLAMWDDTATLTATNIVGCAQFFTKADYPOCOSMHSOILGGTILVIGG 540  
DB 481 LVSGTG YDLCVLAMWDDTATLTATNIVGCAQFFTKADYPOCOSMHSOILGGTILVIGG 540  
QY 541 IIVATLLVFIVILMYRYKVCNHEAPSKMAAASVNVYQNTNGAQPSPSSAPAGAPQGP 600  
DB 541 IIVATLLVFIVILMYRYKVCNHEAPSKMAAASVNVYQNTNGAQPSPSSAPAGAPQGP 600  
QY 601 KVVVRNELLDFATSLARSDSSSSSLGSGEAGLGRAPWRI PPSAPRPKPSLRLMGAF 660  
DB 601 KVVVRNELLDFATSLARSDSSSSSLGSGEAGLGRAPWRI PPSAPRPKPSLRLMGAF 660  
QY 661 ASLDLKSQKELLDSRTAGAGTASARGHSDREPLLGPPAARARSLLPPLGKAKR 720  
DB 661 ASLDLKSQKELLDSRTAGAGTASARGHSDREPLLGPPAARARSLLPPLGKAKR 720

QY 721 SHSFDMDGDFAAAAAGVWPGYSPPRKVSNIWKRSLSVNGMLLPFEESDLVGARGTFGS 780  
DB 721 SHSFDMDGDFAAAAAGVWPGYSPPRKVSNIWKRSLSVNGMLLPFEESDLVGARGTFGS 780  
QY 781 SEWMESTV 789  
DB 781 SEWMESTV 789  
RESULT 2  
AAU28092  
ID AAU28092 standard; Protein; 789 AA.  
XX AAU28092;  
AC AAU28092;  
XX 18-DEC-2001 (first entry)  
XX Novel human secretory protein, Seq ID No 261.  
XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
transgenic animal; Alzheimer's disease; Parkinson's disease; burn;  
amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
gut protection; lung; liver fibrosis; immune deficiency; infection;  
severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
fertility; analgesic; pain; antigen.

OS Homo sapiens.  
XX WO200166689-A2.  
XX 13-SEP-2001.  
XX 05-MAR-2001; 2001WO-US04942.  
XX 07-MAR-2000; 2000US-0519705.  
PR 19-MAY-2000; 2000US-0574454.  
PR 17-JUN-2000; 2000US-0596193.  
PR 14-JUL-2000; 2000US-0616847.  
PR 19-SEP-2000; 2000US-0665363.  
PR 20-OCT-2000; 2000US-0693267.  
XX (HYPSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;  
PI Zhao QA, Yang Y, Dmanac RT, Zhang J, Chen R, Xue AJ, Wang J;  
XX WPI; 2001-589934/66.  
XX N-PSDB; AAS44992.

PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
prepared from various human tissues, for diagnosis and treatment of  
cancer, neurological, inflammatory, and autoimmune disorders -  
XX  
XX Example 4; SEQ ID No 261; 107pp; English.  
XX The invention relates to novel isolated human secreted polypeptides (I)  
and polynucleotides (II). (I) and (II) are useful for treating  
inflammatory conditions such as arthritis, nephritis, Crohn's disease,  
ischaemia-reperfusion injury, shock, sepsis, immune responses, and is  
involved in increasing haematopoiesis, stem cell survival, bone growth  
and remodeling. (I), (II) and modulators of (II) are useful for  
prophylaxis or treatment of one or more cancers. (II) is also useful for  
creating transgenic animals useful for studying the in vivo activities of  
the polypeptide as well as for studying modulators of the polypeptides.  
(I) induces the proliferation of neural cells and regeneration of nerve  
and brain tissue and is useful for the treatment of central and  
peripheral nervous system diseases and neuropathies, such as Alzheimer's,  
Parkinson's disease, Huntington's disease, and amyotrophic lateral  
sclerosis. In addition, (I) is involved in chemotactic or chemokinetic



Accession	Species	Protein	Length	Score	DB	Indels	Gaps
Human	hh00149	protein	SEQ ID NO:4				
Human	ubiquitin-like protein	149Y2H#151	hh00149	brain;			
2-hybrid screening	neuroprotective	signal transducer;					
nervous system disease	diagnosis						
Homo sapiens							
WO2000031255-A1							
02-JUN-2000							
18-NOV-1999							
20-NOV-1998							
(CHUG-)	CHUGAI RES INST MOLECULAR MEDICINE INC.						
Funahashi S, Miyata S;							
WPI: 2000-400066/34							
N-PSDB; AAA60605							
Gene encoding ubiquitin-like protein which interacts with protein							
hh00149, useful in the diagnosis and treatment of diseases associated							
with the nervous system							
Example 2; Page 74-79; 88pp; Japanese							
The present invention describes a ubiquitin-like protein, designated							
149Y2H#151, which interacts with protein hh00149 expressed specifically							
in the brain. The 149Y2H#151 protein has neuroprotective activity, and							
is a signal transducer. The 149Y2H#151 gene and encoded protein are							
useful in the diagnosis and treatment of diseases associated with the							
nervous system. The protein can interact with protein hh00149 expressed							
specifically in the brain. The present sequence represents the human							
hh00149 protein							
Sequence	785 AA;						
Query Match	99.5%; Score 4075; DB 21; Length 785;						
Best Local Similarity	100.0%; Pred. No. 4.9e-302;						
Matches	785; Conservative 0; Mismatches 0; Indels 0; Gaps 0						
QY	5	LGGLAFGMAFAVVDACPKYCVQNLSESLGTLCPSKGLLVFPDPIDRRTVELRGGNFI	64				
Db	1	LGGLAFGMAFAVVDACPKYCVQNLSESLGTLCPSKGLLVFPDPIDRRTVELRGGNFI	60				
QY	65	IHSRQDFANMTGLVDLTLSRNTISHIQPFSLDLESRLSLHDSNRLPSLGEDTLRGLV	124				
Db	61	IHSRQDFANMTGLVDLTLSRNTISHIQPFSLDLESRLSLHDSNRLPSLGEDTLRGLV	120				
QY	125	NLQHLIVNNNOLGGIADAEAFDLTLEDLDSYNNLHGLPWSVRRMVLHQLSLDHLN	184				
Db	121	NLQHLIVNNNOLGGIADAEAFDLTLEDLDSYNNLHGLPWSVRRMVLHQLSLDHLN	180				
QY	185	LDHIAEGTADLOKARLDLTNRLOKLPDPIDFARSQASALTATPFAPPLSFGGNPL	244				
Db	181	LDHIAEGTADLOKARLDLTNRLOKLPDPIDFARSQASALTATPFAPPLSFGGNPL	240				
QY	245	HCNCELLWLRRLERDDDLTETCGSPGKGRYFMWVREEFVCEPPLITQTHKLLVLEQ	304				
Db	241	HCNCELLWLRRLERDDDLTETCGSPGKGRYFMWVREEFVCEPPLITQTHKLLVLEQ	300				
QY	305	AATLKCAIGDPSPLIHWVAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANA	364				
Db	301	AATLKCAIGDPSPLIHWVAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANA	360				
QY	365	AGEATAMVEVSIIVQLPHLSNSTSRTPPKSRSLSDITGSSKTSRGGSGGEGPPKPP	424				
Db	361	AGEATAMVEVSIIVQLPHLSNSTSRTPPKSRSLSDITGSSKTSRGGSGGEGPPKPP	420				
QY	425	AVLVSEVTTTSALVKWSVSKAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFFVNNLVSG	484				
Db	421	AVLVSEVTTTSALVKWSVSKAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFFVNNLVSG	480				
QY	481	LVSGTGYDLCVLAAMDATTATLTATNIVGCAQFTKADYPOCSMHSQILGGTMIIVGG	540				
Db	481	LVSGTGYDLCVLAAMDATTATLTATNIVGCAQFTKADYPOCSMHSQILGGTMIIVGG	540				
QY	541	IIVATLLVFIVILMWRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSAPAGAPQGP	600				
Db	541	IIVATLLVFIVILMWRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSAPAGAPQGP	600				
QY	601	KVVVRNELLDFATSLARASDSSSSSLGSGEAGLGRAPWIPPSAPRPKPSLRLMGAF	660				
Db	601	KVVVRNELLDFATSLARASDSSSSSLGSGEAGLGRAPWIPPSAPRPKPSLRLMGAF	660				
QY	661	ASLDLKSQRKEILLDSRTAGRGAGTSARGHSDREPLIGPPAARSLPLPLEGKAKR	720				
Db	661	ASLDLKSQRKEILLDSRTAGRGAGTSARGHSDREPLIGPPAARSLPLPLEGKAKR	720				
QY	721	SHSFDMGDFAAAAGWPGGYSPRPKVSNIWTKRSLSVNGMLLPFEESDLVGARG					

Db 421 AVLVSEVTTTSLVSKSAPRVKMYQLQNCSDDEVLIYRMIPASNAKAFVNNLVSG 480  
QY 485 TGYDLCLVAMWDDTATLTATNIVGCAQFTTKADYPCQSMHSQLGGTMIIVGIIVA 544  
Db 481 TGYDLCLVAMWDDTATLTATNIVGCAQFTTKADYPCQSMHSQLGGTMIIVGIIVA 540  
QY 545 TLLVFIVILVMRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPTSSAPAGAPPGPKPVV 604  
Db 541 TLLVFIVILVMRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPTSSAPAGAPPGPKPVV 600  
QY 605 RNELLDFTASLASDSSSSSLGSCFAAGLGRAPRIPPSAPRPXPSIDRLMGAFASLD 664  
Db 601 RNELLDFTASLASDSSSSSLGSCFAAGLGRAPRIPPSAPRPXPSIDRLMGAFASLD 660  
QY 665 LKSQRKEELDSRTAGRCAGTSARGHSDREPLGPPAARSLPLPLEGKAKRSHSF 724  
Db 661 LKSQRKEELDSRTAGRCAGTSARGHSDREPLGPPAARSLPLPLEGKAKRSHSF 720  
QY 725 DMGDFAAAAAGGVVPGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFSSEWV 784  
Db 721 DMGDFAAAAAGGVVPGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFSSEWV 780  
QY 785 MESTV 789  
Db 781 MESTV 785

RESULT 5  
AAG67505  
ID AAG67505 standard; Protein; 766 AA.  
XX AC AAG67505;  
XX DT 26-NOV-2001 (first entry)  
XX DE Amino acid sequence of a human secreted polypeptide.  
XX KW Human; secreted polypeptide; nervous disease; muscular disease; tumour;  
KW gastrointestinal ulceration; spinal cord disease; trachea disease;  
KW thyroid gland disease; ovary disease; prostate disease; heart disease;  
KW renal gland disease; small intestine disease; thymus disease;  
KW lymph node disease; muscular system disease; colon disease;  
KW lipase deficiency; cystic fibrosis; pancreatitis; clot formation;  
KW myocardial infarction; angioplasty; liver disease; coagulation disorder;  
KW microbial disease; immune disorder; inflammation; transplant rejection;  
KW bone thickness; bone density; ferroxidase loss; apoptosis;  
KW vascular smooth cell proliferation; vaccine.  
XX OS Homo sapiens.  
XX PN WO200166690-A2.  
XX PD 13-SEP-2001.  
XX PF 05-MAR-2001; 2001WO-US07143.  
XX PR 06-MAR-2000; 2000US-0187107.  
PR 13-MAR-2000; 2000US-0188916.  
PR 03-OCT-2000; 2000US-0236874.  
PR 03-OCT-2000; 2000US-0237846.  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;  
XX WPI; 2001-570768/64.  
DR N-PSDB; AAH78197.  
XX Novel isolated secreted polypeptide useful for treating nervous and  
PT muscular diseases, gastrointestinal ulceration, coagulation and immune  
PT disorders, microbial diseases, inflammation and transplant rejection -

XX Claim 1; Page 58-60; 102pp; English.  
XX The present sequence represents a human secreted polypeptide. The  
CC secreted polypeptides and polynucleotides are useful for treating  
CC nervous and muscular diseases, for inhibiting tumour formation and  
CC metastasis, for treating gastrointestinal ulceration, for preventing  
CC and treating diseases in spinal cord, thyroid gland, ovary, prostate,  
CC renal gland, small intestine, heart, trachea, thymus, lymph node,  
CC muscular system and colon, for treating lipase deficiency in cystic  
CC fibrosis and pancreatitis, for treating undesirable clot formation  
CC such as myocardial infarction, during angioplasty and all surgical  
CC procedures that require decreased blood clot formation, for treating  
CC liver diseases, coagulation disorders and microbial diseases, for  
CC treating immune disorders, for treating inflammation and transplant  
CC rejection, for enhancing bone thickness and increasing bone density,  
CC for reducing the loss of essential ferroxidases, for suppressing  
CC apoptosis, and for regulating vascular smooth cell proliferation. They  
CC may also be used as vaccines.  
XX SQ Sequence 766 AA;  
Query Match 45.1%; Score 1846; DB 22; Length 766;  
Best Local Similarity 48.5%; Pred. No. 5.3e-132;  
Matches 383; Conservative 128; Mismatches 225; Indels 54; Gaps 11;  
QY 1 METLLGGLAFGMFAVVDACPKYVCQNLSESLGTLCPKSGLLFPVPPIDRRTELRLG 60  
Db 1 MEKVLFLFLIGIA-VKAIQCPKRCVCQLLSPNLATLCAKKGLLFPVPPIDRRTELRLA 59  
QY 61 GNFIHISQDPANMTGLVDLTLSRNTISHIOPFFSLDLESRLSHLDNRNLPUGEDTL 120  
Db 60 DNFVTNIKKDFANMTSLVDLTLSRNTISFITPHAFADLRNLRALHLSNRLTKINDMF 119  
QY 121 RGLVNLHLIYVNNOLGGIADEFDLTLLEDDLSYNNLHCLPWSVRRVNNHOLSL 180  
Db 120 SGLSNLHLILNNOQLTSLISSTAFDD-VFALBELDLSYNNLETIPMDAVEKWSLHTLSL 178  
QY 181 DHNLHDHIAEGTFADLQKLARLDLTSLNRLOKLPPOPIFARSOASALTATFPAPPLSFG 240  
Db 179 DHNMIDNIPKGFTHLHKMTLDVTSNKLQKLPDLPFORAQVLAISGIIISPSTFALSFG 238  
QY 241 GNPLHCNCELLMIRLERDDDETCSGPGGLKGRVFWHVREBEFVCEPPLIIOHTHKLIV 300  
Db 239 GNPLHCNCELLMIRLRSREDDLETSCASPPLLTGRYFWSIPEBEFLCEPPLIIRHTEMRV 298  
QY 301 LEGQATLKCKAIGDPSPLIHVAPDDRLVGNSSRTAVYDNGTLDIFITSDSGAFTCI 360  
Db 299 LEGQATLRCRKARGDPEPAIHWSIPEGLKLSNATRSVVDYDNGTLDILITTVKDTGFTCI 358  
QY 361 AANAAGEATAMVEVSTVQLPHLSNSTRTAPPKRSLSITGSSKTSRGGGGGGGEPKPS 420  
Db 359 ASNPAGEATQIVDLHIKLPNLLNSTNHIHEPDGSSDLSITKSGNSTSSNG---DTK 415  
QY 421 PPERAVLVSEVTTTSLVSKSAPRVKMYQLQNCSDDEVLIYRMIPASNAKAFVNN 480  
Db 416 LSQDKIVVAEATSSALLKFNFORNIPGIRMEQIQVNGTYDDTLVYRMIPPTSKTELNN 475  
QY 481 LVSGTGYDLCLVAMWDDTATLTATNIVGCAQFTTKADYPCQSMHSQLGGTMIIVG 540  
Db 476 LAAGMTYDLCLVLAIVDDGITSATATRVWGCIOFTTEQDYVRFCHFMOSQLGGTMIIVG 535  
QY 541 IIVATLLVFIVILVMRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPTSSAPAGAPPG 600  
Db 536 IIVASVLVFIILMIRYKVCNNNGOHKV-TKVSNNVYSQTNGAQ-----IOGCVTLFQSVS 590  
QY 601 KVVVRNELDFTASLARASDSS---SSSLGSGEAGLGRAPRIPPS----- 645  
Db 591 KQAVGHE---ENACCKATSDNVIQSSSETCSSODSSTTTSA---LPPTSSTSVSQKQK 644  
QY 646 -APRPKPSLDRLMGAFASLDLKSQKEE-----LLDSRTAGRCAGTSARGHSDREPLIG 700  
Db 645 RKTGTPKEPQNEAVTNVESQNTNRNNTALQLASRPDPSVTEGTSKRAHIKPSKFI 704

QY 701 PPAARARSLLPLEGKAKRSHSFDMGDFAAAAAGVPGGYSPPRKVSNIWTKBSLSVN 760  
Db 705 LPAERSGARHKYSLNGELKEYYC-----XINSPNTCGLFPKRSMSMN 746  
QY 761 GMLLPPEESD 770  
Db 747 VMFIQSDCSD 756

RESULT 6

ABP43533  
ID ABP43533 standard; Protein; 719 AA.

XX AC ABP43533;

XX DT 08-AUG-2002 (first entry)

XX DE Human secreted protein (SCEP) 57.

XX KW Human; secreted protein; SECP; SECP expression; gene therapy;  
KW protein therapy; immune system disorders; AIDS; thymic hypoplasia;  
KW anaemia; asthma; Crohn's disease; neurological disorder; epilepsy;  
KW Huntington's disease; dementia; Parkinson's disease; Down's syndrome;  
KW developmental disorder; cell proliferative disorder; cancer.

XX OS Homo sapiens.

XX XX WO200226982-A2.

XX XX 04-APR-2002.

XX XX 25-SEP-2001; 2001WO-US30042.

XX XX 29-SEP-2000; 2000US-236869P.

XX XX 11-OCT-2000; 2000US-239812P.

XX XX 12-OCT-2000; 2000US-240108P.

XX XX 17-OCT-2000; 2000US-241282P.

XX XX 20-OCT-2000; 2000US-242218P.

XX XX (INCY-) INCYTE GENOMICS INC.

XX PI Yue H., Tang YT, Nguyen DB, Yao MG, Xu Y, Tribouley CM;  
PI Sanjanwala MS, Walla NK, Baughn MR, Sapperstein SK, Lal P;  
PI Thornton M, Gandhi AR, Ramkumar J, Elliott VS, Arvizu C;  
PI Thangavelu K, Gietzen KJ, Ding L, Au-young J, Tran B, Policky JL;  
PI Lee S, Lu DAM, Burford N, Warren BA, Gururajan R, Duggan BM;  
PI Honchell CD, Hafalia AJA;

XX WPI; 2002-394239/42.

XX N-PSDB; ABN99416.

XX PT New human secreted proteins, useful for diagnosing, treating or  
PT preventing immune system disorders (e.g. Crohn's disease), neurological  
PT disorders (e.g. Parkinson's disease), or cell proliferative disorders  
PT (e.g. cancers).

XX PS Claim 1; Page 188-199; 238pp; English.

XX The invention comprises the amino acid and coding sequences of human  
CC secreted proteins (SECP). The SECP DNA and amino acid sequences of the  
CC invention are useful for treating/preventing disorders associated with  
CC decreased or elevated expression of SECP. The SECP DNA and protein  
CC sequences are specifically useful for treating/preventing (i.e. gene  
CC therapy and protein therapy): immune system disorders (e.g. AIDS, thymic  
CC hypoplasia, anaemia, asthma or Crohn's disease); neurological disorders  
CC (e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease);  
CC developmental disorders (e.g. Down's syndrome); and cell proliferative  
CC disorders (e.g. cancer). The proteins ABP43477 - ABP43543 represent the  
CC human secreted proteins (SECP) of the invention.

XX SQ Sequence 719 AA;

Query Match 44.5%; Score 1820.5; DB 23; Length 719;  
Best Local Similarity 52.1%; Pred. No. 4.3e-130;  
Matches 371; Conservative 114; Mismatches 190; Indels 37; Gaps 11;  
QY 1 METILGLLAFGMAFAVVDACPKYCVCONLSESLGTLCPKGLLVDPDIDRRIVELRLG 60  
Db 1 MEKILFYLFLIGIA-VKAIQCPKRCVCQILSPNIATLCAKGLLVFPDIDRRIVELRLA 59  
QY 61 GNFIHISRODPANMTGLVLTLSRNTISHIQPFSLDLSRLSLHLSDNLSPLGSDTL 120  
Db 60 DNFVNIKKDFANMTSLVLTLSRNTISFITPHAFADLRNLRALHLNSRLTKITDMF 119  
QY 121 RGLVNLQHLIYNNOLGGIADAEFLLTLEDLDSYNNLHGLPWDSVRRMNLHQLSL 180  
Db 120 SGLSNLHLILNNQLTLISSTAFDD-VFALEELDLSYNNLETIPWDAVEKMSLHTLSL 178  
QY 181 DHNLIDHTAEGTFADLOKLARLDLTNSRLOKLPDPIFARSOASALTATPEAPPLSFSG 240  
Db 179 DHNMIDNIPKGTFSHLHKWTRLDVTSNKLQKLPDPLFORAQLVATSGIISFSTFALSFG 238  
QY 241 GNPLHCNCELLWRLRLERDDLETCGSPGLKGRYFMHVREEFVCEPLITQHTHKLVL 300  
Db 239 GNPLHCNCELLWRLRLERDDLETCASPLLTGRYFWSIPPEEFLCEPLITRHTHEMRV 298  
QY 301 LEGQAATLKCAIGDPSPLIHWAPDDRLVGNSSRTAYDNGTLDIFITTSQDSGACTCI 360  
Db 299 LEGQRAATLRCKARGDPEPAIHWISPEGLISNATRSVYDNGTLDILITTVKDTGACTCI 358  
QY 361 AANAAGEATAMVEVSIVOLPHLSNSTSRTAPPKSRSLDITGSSKTSRGGSGGGEPPKS 420  
Db 359 ASNPAGEATQIVDLHLIIKLPHLLNSTNIHEPDPGSSDITSTKSGSNSSNG---DTK 415  
QY 421 PPERAVLVSEVTTTSALVKWSVKSAPRVKMYQLQYNCSDDEVLIRMI PASNKAFVNN 480  
Db 416 LSQDKIVVAEATSSALLKFNFORNIPGRWFIQYNGTYDDTLVYRMIPTTSKTLVNN 475  
QY 481 LVSGTGYDLCLVAMWDDTATLTATNIVGCAQFFKADYPCQSQMHSQILGCTMILVIGG 540  
Db 476 LAAATMYDCLVLAIFYDDGITSLTATRVVGCIOFTTEQDYVRCHFQSQOPLGGMIIIGG 535  
QY 541 IIVATLLVFIIVLMVRYKVCNHEAPSKMAAAVSNVYVSTNGAQPPPPPSAPAGAPQGGP 600  
Db 536 IIVASVLFVLIIMIRYKVCNNNGHVK-TKVSNNVYVSTNGAQ-----IQGCVTLPPQSVS 590  
QY 601 KVVVRNELDFTASLARASDSS---SSSLGSGEAGLGRAPWRIPPS-----645  
Db 591 KQAVGHE--ENAOCCKATSDNVIQSSSETCSSQDSSTTTSA---LPPSWTSTSVSQKQ 644  
QY 646 -APRPKPSLDRLMGAFASLDLKSQKEE---LLDSRTPAGRGAG-TSARGH 691  
Db 645 RKTGTPKSTEPQNEAVTNVESQNTNRNNTALQLASRPDPSVTEGPTSKRAH 696

RESULT 7

AAG67512  
ID AAG67512 standard; Protein; 771 AA.

XX AC AAG67512;

XX DT 26-NOV-2001 (first entry)

XX DE Amino acid sequence of a human secreted polypeptide.

XX KW Human; secreted polypeptide; nervous disease; muscular disease; tumour;  
KW gastrointestinal ulceration; spinal cord disease; trachea disease;  
KW thyroid gland disease; ovary disease; prostate disease; heart disease;  
KW renal gland disease; small intestine disease; thymus disease;  
KW lymph node disease; muscular system disease; colon disease;  
KW lipase deficiency; cystic fibrosis; pancreatitis; clot formation;  
KW myocardial infarction; angioplasty; liver disease; coagulation disorder;  
KW microbial disease; immune disorder; inflammation; transplant rejection;  
KW bone thickness; bone density; ferroxidase loss; apoptosis;  
KW vascular smooth cell proliferation; vaccine.





[illegible]





	The present sequence represents a human secreted polypeptide. The
	secreted polypeptides and polynucleotides are useful for treating
	nervous and muscular diseases, for inhibiting tumour formation and
	metastasis, for treating gastrointestinal ulceration, for preventing
	treatng diseases in spinal cord, thyroid gland, ovary, prostate,
	renal gland, small intestine, heart, trachea, thymus, lymph node,
	mucular system and colon, for treating lipase deficiency in cystic
	fibrosis and pancreatitis, for treating undesirable clot formation
	such as myocardial infarction, during angioplasty and all surgical
	procedures that require decreased blood clot formation, for treating
	liver diseases, coagulation disorders and microbial diseases, for
	treating immune disorders, for treating inflammation and transplant
	rejection, for enhancing bone thickness and increasing bone density,
	for reducing the loss of essential ferroxidases, for suppressing
	CC apoptosis, and for regulating vascular smooth cell proliferation.
	CC They may also be used as vaccines. XX XX
SQ	Sequence           628 AA;
	Query Match         38.4%; Score 1573; DB 22; Length 628; Best Local Similarity 53.1%; Pred. No. 2.7e-111; Matches 312; Conservative 88; Mismatches 174; Indels 14; Gaps 5
QY	16 AVVDACPKYCVCQNLSBSLGTLCPSKGLLFVPPDIDRTVELRGCGNFIHISRODFAMM 75   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
D.	23 ATPSPCPRRCRCQTQSPLSVLPFCAGLLFVPSPDLDRRAELRADNFIASVRERDLANM 82   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY
Y	76 TGLVDLTLSRNTISHIQPFPSFLDESRLSHLDNSRLPSLGEDTLURGIUVNLQHILVNNO 135   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
B	83 TGLHLHSLRNTIRHVAAAFADRALRALHDGNRLTSLGEGQGUGIUNRHULTSNNQ 142   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY
Y	136 LGGIADFATEDFLTLEDLDSLYNNHLPMDSVRRMVNLHQLSLDHNLLDHIAEGETAD 195   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
B	143 LAALAAGALDDCAETLEDDLDSLYNNLEQLPWEALGRLGNVNTIGLDHNILLASVPAGEFR 202   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY
Y	196 LOKLARIDLTSNRLOKUPODPPIPARSQASALTATFPAPPLS---PFGNGNPJHCNCCELLW 252   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
B	203 LHKLARLDMTSNRUUTTTPPDPLFSR--LPLLARPGRGPSAALVLAFGGNPLHCNCCELWW 259   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY
Y	253 LRRLERDDDETCSGPGCLKRGYPFWHVREBEFVCEPPLITOTHTKKLVLEGQAATLKCKA 312   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
B	260 LRRLARECDDELAECASPALGGRYFWAVGEBEFVCEPVVTUHRPPPLAVPAGRPAALRCRA 319   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY
Y	313 IGDPSLIHWVAPPDRDLVGNSSRTAVDYONGTLDIPITTSDSGAFTCIAANAAGEATAMV 372   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
B	320 VGDPEPRVMWSPCQRLLGNSSRARAFPNGTILELVTEPGDGIFTCIAANAAGEATAAV 379   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY
Y	373 EVSI--VOLPHLSNSTSTAPPKGRLSDI TGSSKTSTRGGSGGS GGEPKPSPPERAVLVESE 430   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
B	380 ELTVGPPPPPQOLANSSTDOPRDGPDPDALTFPPSAASA SAVDTGC---PTPTRGVQYTE 435   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY
Y	431 VTFTTSALKVSYSKSARKVKMYOQYNCSDDDEVILIYRMIPASNKAFAVVNNLVSGTYDLC 490   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
B	436 HGATAALVQPMDQRPICGRIWMYOQYNSADDILVIYRMIPAESRFILTDLASGRITYDLC 495   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY
Y	491 VLAWMDTOTATTLATNIIVGCAQFFT KADYPOCOSMH SQII LGCTMTILVII GGIIVATLLVFI 550   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
B	496 VLAYVEDSATGLTATRPGCARFSTEPALRPCGAHPFLGCTMIIALGGVIVASVLVFI 555   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY
Y	551 VILMVRKYKCNEHA P--SKMAAAA VN NVSY QTN GA O P P P S S A GAP PP 596   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : DB.
B	556 FVLLMRKVYHGQQPGAKAPIAPVSSVSCQTN GAL GP TP T PA P AP E P 603   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
RESULT 13	
AAG65805	ID
xx	AAG65805 standard; Protein; 628 AA.
xx	AAG65805;
xx	(first entry)
xx	Human leucine-rich repeat (LRR) family member, 33995 polypeptide.

Best Local Similarity 52.9%; Pred. No. 9.2e-111;  
Matches 311; Conservative 88; Mismatches 175; Indels 14; Gaps 5;  
QY 16 AVVDACPKYCVQNLSESLGTLCPKSGLLFVPPDIDRTVELRLGNGFIHISRQDFANM 75  
Db 23 ATPSPCPRRCRCQTSPLSVLCPGAGLLFVPPSIDRRAELRLADNFASVRRDLANM 82  
QY 76 TGLVDTLSRNTISHIQPFSLDLSRLSHLSDNRLPSLGEDTLRGLVNLQHLIVNNQ 135  
Db 83 TGLHLSLSRNTIRHVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSNNQ 142  
QY 136 LGGIADEAFEDFLITLEDLDSYNNHLGLPWSVRRMVLHQLSDHNLHDHIAEGTFAD 195  
Db 143 LAALAAGALDDCAETLEDLDSYNNLEQLPWEALGRGLGNVTLGLDHLNLLASVPGAFSR 202  
QY 196 LOKLARLDLTSNRLOKLPDPFARSQASALTATPAPPLS---FSGGNPLHNCCELLW 252  
Db 203 LHKLARLDLTSNRLTTPDPLFSR---LPLLARPRGSPASALVLAFGNPLHNCCELLW 259  
QY 253 LRLERDDLETGSGGGLKGRYFWHVRBEEFVCEPPLITQHTKLLVLEGOAATLKCKA 312  
Db 260 LRLAREDDLEACASPPALGGRYFWAVGEEFVCEPVPVTHRSPLAVPAGRPALURCKA 319  
QY 313 IGDPSLLIHWVADDDRLVGNSSRTAVYDNGTLDIFITTSQDGAFTCIANAAGATAMV 372  
Db 320 VGDPEPRVRVWSQGRLLGNSSRARAPNGTLELLVTEPGDGGIFTCIANAAGATAAV 379  
QY 373 EVSI--VQLPHLSNSTRTAPPKSRSLSDITGSKTSRGGGSGGPPKSPERAVLUSE 430  
Db 380 ELIVGPPPPQLANSTCDPPRGGDPAITPPSAASAKVADTG---PPTDRGVQVTE 435  
QY 431 VTTTSALVKSWSKAPRVKMYQLQVNCSDDEVLIYRMTPASNKAFVWNNLVSGTGYDLC 490  
Db 436 HGATAALVQWPDQRPPIGIRMYQIQVNSSADDILVYRMIPAESRSLTDLASGRYDLC 495  
QY 491 VLAWDDTATLTATNTVGCAPQFTKADYPQCOSMHSQILGGTMILVIGGIIVATLLVFI 550  
Db 496 VLAVGESATGLTATRVGCARSTEPALPCGAPHAPFLGGTMIIALGGVIVASVLVFI 555  
QY 551 VILMRYKVCNHEAP--SKMAAAVSNVYVYOTNGAQPSPPSAGAPAPP 596  
Db 556 FVLLMRYKHGGQPPGKAKIPAPVSSVCSOTNGALGTPPTAPPAPPEP 603  
RESULT 14  
ID ABG34079  
AC ABG34079  
XX  
XX  
DT 15-JUL-2002 (first entry)  
DE Human Pro peptide #51.  
XX  
XX Human; PRO; secreted protein; transmembrane protein;  
KW Genetic disorder; tumour; cancer.  
XX  
XX Homo sapiens.  
OS WO200224888-A2.  
XX  
XX  
PN 28-MAR-2002.  
XX  
PD 29-AUG-2001; 2001WO-US27099.  
XX  
XX 01-SEP-2000; 2000US-229896P.  
PR 05-SEP-2000; 2000US-230621P.  
PR 22-SEP-2000; 2000US-235147P.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 12-JAN-2001; 2001US-261878P.  
PR 16-JAN-2001; 2001US-261910P.  
PR 16-JAN-2001; 2001US-261939P.  
PR 16-JAN-2001; 2001US-262150P.

25-JAN-2001; 2001US-264395P.  
02-FEB-2001; 2001US-266421P.  
09-FEB-2001; 2001US-267623P.  
28-FEB-2001; 2001WO-US06520.  
03-MAR-2001; 2001US-274399P.  
04-APR-2001; 2001US-280982P.  
04-APR-2001; 2001US-282129P.  
04-APR-2001; 2001US-282199P.  
09-MAY-2001; 2001US-290589P.  
25-MAY-2001; 2001WO-US17092.  
01-JUN-2001; 2001WO-US17800.  
20-JUN-2001; 2001WO-US19692.  
29-JUN-2001; 2001WO-US21066.  
09-JUL-2001; 2001WO-US21735.  
(GETH ) GENENTECH INC.  
Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,  
Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;  
Fong S;  
WPI; 2002-362426/39.  
N-PSDB; ABK70011.  
New PRO polypeptides and polynucleotides encoding the polypeptides.  
useful in gene therapy, chromosome identification, tissue typing, or  
for genetic analysis of individuals with genetic disorders  
Claim 11; Figure 102; 218pp; English.  
This invention relates to the cDNA and protein sequences of novel  
secreted and transmembrane polypeptides PRO polypeptides. The  
invention also comprises a method for producing the proteins of the  
invention by recombinant means and antibodies specific for the protein  
of the invention. The antibody may be used for detecting the PRO  
proteins of the invention and may be used to modify their activity.  
Polynucleotides may be used as hybridisation probes for a cDNA library  
to isolate the full-length PRO cDNA or to isolate other cDNAs, to  
construct hybridisation probes for mapping the gene which encodes that  
PRO and for genetic analysis of individuals with genetic disorders, in  
assays to identify other proteins or molecules involved in binding  
reaction, to generate transgenic animals or knock-out animals which in  
turn are useful in the development and screening of therapeutically  
useful reagents, for chromosome identification, and tissue typing. The  
PRO polypeptides are useful in gene therapy, and as molecular weight  
markers for protein electrophoresis purposes. The sequences may  
also be used to detect overexpression on PRO polypeptides in cancerous  
tumours and for screening for differentially expressed genes using  
microarray technology. The present sequence represents a human PRO  
protein of the invention.  
Sequence 627 AA  
Query Match 38.2%; Score 1563.5; DB 23; Length 627;  
Best Local Similarity 53.1%; Pred. No. 1.4e-110;  
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;  
QY 16 AVVDACPKYCVQNLSESLGTLCPKSGLLFVPPDIDRTVELRLGNGFIHISRQDFANM 75  
Db 23 ATPSPCPRRCRCQTSPLSVLCPGAGLLFVPPSIDRRAELRLADNFASVRRDLANM 82  
QY 76 TGLVDTLSRNTISHIQPFSLDLSRLSHLSDNRLPSLGEDTLRGLVNLQHLIVNNQ 135  
Db 83 TGLHLSLSRNTIRHVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSNNQ 142  
QY 136 LGGIADEAFEDFLITLEDLDSYNNHLGLPWSVRRMVLHQLSDHNLHDHIAEGTFAD 195  
Db 143 LAALAAGALDDCAETLEDLDSYNNLEQLPWEALGRGLGNVTLGLDHLNLLASVPGAFSR 202  
QY 196 LOKLARLDLTSNRLOKLPDPFARSQASALTATPAPPLS---FSGGNPLHNCCELLW 252  
Db 202 LHKLARLDLTSNRLTTPDPLFSR---LPLLARPRGSPASALVLAFGNPLHNCCELLW 258







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 21:27:43 ; Search time 36 Seconds  
(without alignments)  
644.852 Million cell updates/sec

Title: US-09-831-846-2  
Perfect score: 4094  
Sequence: 1 METLLGGLAFGMAFAVDA.....DLVGARTGSGSEWNSTV 789

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/2/aaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/aaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/aaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/aaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/aaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/2/aaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	368	9.0	708	US-09-131-648-2	Sequence 2, Appli
2	368	9.0	1091	US-08-986-485-5	Sequence 5, Appli
3	351	8.6	1101	US-08-986-485-2	Sequence 2, Appli
4	299	7.3	673	US-09-063-950-2	Sequence 2, Appli
5	284.5	6.9	1525	US-09-191-647-2	Sequence 2, Appli
6	284.5	6.9	1525	US-09-540-245A-2	Sequence 2, Appli
7	284.5	6.9	1525	US-09-540-153-2	Sequence 2, Appli
8	268	6.5	1523	US-09-182-024A-2	Sequence 2, Appli
9	263	6.4	1480	US-09-191-647-7	Sequence 7, Appli
10	263	6.4	1480	US-09-540-245A-7	Sequence 7, Appli
11	263	6.4	1480	US-09-540-153-7	Sequence 7, Appli
12	263	6.4	1480	US-09-182-024A-5	Sequence 5, Appli
13	263	6.4	1480	PCT-US91-09055-2	Sequence 2, Appli
14	260.5	6.4	222	PCT-US91-09055-3	Sequence 3, Appli
15	259	6.3	231	US-08-986-485-7	Sequence 7, Appli
16	248	6.1	603	US-08-190-802A-50	Sequence 50, Appl
17	248	6.1	603	US-08-477-346-50	Sequence 50, Appl
18	248	6.1	603	US-08-473-089-50	Sequence 50, Appl
19	248	6.1	603	US-08-487-072A-50	Sequence 50, Appl
20	242.5	5.9	353	US-08-986-485-6	Sequence 6, Appli
21	241.5	5.9	605	US-08-190-802A-49	Sequence 49, Appl
22	241.5	5.9	605	US-08-477-346-49	Sequence 49, Appl
23	241.5	5.9	605	US-08-473-089-49	Sequence 49, Appl
24	241.5	5.9	605	US-08-487-072A-49	Sequence 49, Appl
25	236.5	5.8	649	US-09-188-930-305	Sequence 305, App
26	235.5	5.8	605	US-09-063-950-5	Sequence 5, Appli
27	224	5.5	373	US-09-724-864-43	Sequence 43, Appl

28	222	5.4	180	3	US-08-986-485-8	Sequence 8, Appli
29	222	5.4	224	5	PCT-US91-09055-4	Sequence 4, Appli
30	203	5.0	821	1	US-08-339-578-2	Sequence 2, Appli
31	201	4.9	320	1	US-07-613-083B-1	Sequence 1, Appli
32	198	4.8	610	1	US-07-821-717B-6	Sequence 6, Appli
33	198	4.8	610	1	US-08-119-262B-6	Sequence 6, Appli
34	198	4.8	610	1	US-08-135-923A-11	Sequence 11, Appl
35	198	4.8	610	1	US-08-234-265A-11	Sequence 11, Appl
36	197	4.8	560	3	US-08-592-500-2	Sequence 2, Appli
37	197	4.8	560	3	US-08-195-006-2	Sequence 2, Appli
38	197	4.8	560	4	US-09-063-950-4	Sequence 4, Appli
39	197	4.8	560	5	PCT-US94-07644A-2	Sequence 2, Appli
40	189	4.6	434	4	US-09-540-245A-19	Sequence 19, Appl
41	186.5	4.6	375	1	US-08-303-238-2	Sequence 2, Appli
42	186.5	4.6	375	1	US-08-458-834-2	Sequence 2, Appli
43	185	4.5	612	2	US-08-359-705B-8	Sequence 8, Appli
44	185	4.5	612	2	US-08-286-846A-8	Sequence 8, Appli
45	185	4.5	612	2	US-08-457-880A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1  
US-09-131-648-2  
; Sequence 2, Application US/09131648  
; Patent No. 6168920  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Patterson, Chandra  
; TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS  
; FILE REFERENCE: PF-0576 US  
; CURRENT APPLICATION NUMBER: US/09/131,648  
; CURRENT FILING DATE: 1998-08-10  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 2  
; LENGTH: 708  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: 2687731  
US-09-131-648-2

Query Match	9.0%	Score 368;	DB 4;	Length 708;
Best Local Similarity	22.6%	Pred. No. 3.4e-22;		
Matches 154;	Conservative 98;	Mismatches 238;	Indels 192;	Gaps 27;
QY	1	METLLGGLAFGMAFAVDA---CPKYVCQ-----NLSESLGTL-CPKGLLVFP 48		
Db	8	IVLVLG--LAITTLVQAVDKKVDCCPLCTCEIRPMTPRSIYMEASTVDCNDLGLTTPFA 65		
QY	49	DIDRTEVLRLGGNFIIHIS-RQDF-ANMTG-----EDFLTLEDLD-LSY-- 158		
Db	66	REPANTQILLQTNIAKEYTEYDFPVNLTGLDLSQNNLSSVTNINVKMPQLLSVYLEE 125		
QY	78	-----LVDLTLSRNTISHIQFSPFLDLSRLSLHLDNSRFLSGEDTURL 121		
Db	126	NKLTPEKCLSELNQLQELYNHNLSTISPGAFGLHNLRLHLNLSRNLQMINSKWFD 185		
QY	122	GLVNLQHLIVNNQLGGIADEAF-----EDFLTLEDLD-LSY-- 158		
Db	186	ALPNLEILMIGENPIIRIKDMNFKPLNLSRVLVAGINLTPEIDNALVGLNDEISFYD 245		
QY	159	NNLHGLPWSVRMVLNLSLHLDLHIAEGTFA----- 194		
Db	246	NRLIKVPHVALQKVNLFELNKPINRIIRGDFSNMLHLKELGINNMPELISIDSLAV 305		
QY	195	-DQKLARDLTSN-RLQKLPPDPITFARQASALTATPFA-----PPL-SFSF 239		

Db 306 DNLPLDKTEATNPRLSYIHPNAPFRLPKLESMLNSALSYHGTIESLNLKEISI 365  
Qy 240 GGNPLHCNCELWLW-----RLRDDDLTETCGSPGLKGRYFWHVRREEF-----VCEP 288  
Db 366 HSNPIRCDVIRMMNMKNIRIFMEPDSL-FCVDPEFQOQ---NVRQVHFRDMMIEICLP 421  
Qy 289 PLITQHTHLLVLE-GQAATLKCAIGDPSPLIHWAPD-DRLVGN--SSRTAVYDNGTL 344  
Db 422 LIAPESFPNINVEAGSYVSFICRATAEPQELIWIITPSGQKLLPNTLTDKIFYVHSEGL 481  
Qy 345 DIFITTSQSGAFTCIAANAAGEATAAMVEVSIVQLPHLSNSTSRTPAPPKSRSLSDITGSSK 404  
Db 482 DINGVTPREGGLVTCIATNLVG-----ADLSVMIKIKVDGSP 518  
Qy 405 TSGGGGSGGPPKPPPERAVLSEVTTTSALVKSVS-----KSAPRVKMYQLQVNCSD 460  
Db 519 QDNNGSLN-----IKIRDIQANSVLVSWKASSKILKSSVKWTAF---VKTEN 562  
Qy 461 DEVLIYRMI PASNKA FVNNLVSGTGYDLCVLAMWDDTATLTATNIVGCAQFFTKADYP 520  
Db 563 SHAAQASARIPSDVKVNNLTHLPSTBYKICI-----DIPTIYQKNRKKCVNVTYKGLHP 616  
Qy 521 QCQSMHSQILGTMILVIGII 542  
Db 617 D-QKEYEKNNTTLMACLOGLL 637

## RESULT 2

US-08-986-485-5  
; Sequence 5, Application US/08986485  
; Patent No. 6046030  
; GENERAL INFORMATION:  
; APPLICANT: WU, SHUIJIAN  
; APPLICANT: SWEET, RAYMOND  
; APPLICANT: TRUNEH, ALEMSEGED  
; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RATNER & PRESTIA  
; STREET: P.O. BOX 980  
; CITY: VALLEY FORGE  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19482  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: .Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/986,485  
; FILING DATE: 08-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/059,448  
; FILING DATE: 22-SEP-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PRESTIA, PAUL F  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GH-70264  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1091 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-986-485-5

Query Match 9.0%; Score 368; DB 3; Length 1091;  
Best Local Similarity 21.3%; Pred. No. 6.6e-22;  
Matches 205; Conservative 92; Mismatches 326; Indels 340; Gaps 33;  
Qy 30 LSESLGTLCPSKGLL-----FVPPDIDRTVELRLCGNFIHISRODFANMTGLVDT 82  
Db 188 LSRSLTLRLSKNRITQLPVKAFKL-----RLTQLDLNRNRIIEGTFTQGLDSLEVL 243  
Qy 83 LSRNTISHIOPPSFLDLSRLSLHLSNRLPSLGEDTLRLGLVNLQHLIWNNOGLGIADE 142  
Db 244 LQRNNISLTGCAFGLSKMHLVHLEYNLSLVNNGSLYGLTALHQLHLSNNSISRIORD 303  
Qy 143 AFEDFLTLLEDLSYNNLHGLPMDSVRRMVNHLSDLDHLLDIAEGTTFADLOKLARL 202  
Db 304 GW-SFCQKLHELILSFNNLTRLDEESLAELSSLSILRLSHNAISHIAEGAFKGLSLRVL 362  
Qy 203 DLTSNRLQKLPDPIPARSOASALT-ATPPAPPLS-----FSFGGNPLH- 245  
Db 363 DLDHNEISGTTIEDTSGAFTGLDNLKSLTLPGNKIKSVAKRAPSGLESLEHLNLGNAIRS 422  
Qy 246 -----CNCELLWL-----RRLERDDDLTETCGSPGLKGRYFW 277  
Db 423 VOFDAFAKMKLKYIYSSSEFLDCQKWLPPWLMGRMLQAFVTATCAHPESLKGQSF 482  
Qy 278 HVREEFYCE-----PPLITQ----- 293  
Db 483 SVLPDSFVCDOPPKPQIITQPETTMAVVGKDIRFTCSAASSSSSPMTFAWKDNEVLANA 542  
Qy 294 -----HTHK----- 297  
Db 543 DMENFAHRAQDGEVMEYTTILHLRHVTFHGEGRYQCIITNHFSTYSHKARLTNVNLP 602  
Qy 298 -----LLVLEGOAATLKCAIGDPSPLIHW-----VAPDDRVLGNS 333  
Db 603 FTKIPIHDIAIRGTTRALECATGHPNFOIAWKQDGGTDFPAARERRMHVMDDD----- 656  
Qy 334 SRTAVYDNGTLDIFITTS-----ODSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTPA 390  
Db 657 -----DVFFITDVKIDDMGVYSCTAQSAGSVSANATLVLETPSL-----A 698  
Qy 391 PPKSLSLDITGSSKTSRGGGSGGEPKSPPERAVLYSEVTTTSALVKSYSKSAAPRYK 450  
Db 699 VPELE-----DRVVTVGE-----TVAFQCKATGSPTPRIT 727  
Qy 451 MYQLQVNGSDDEVLIYRMI PASNKA FVNNLVSGTGYDLCVLAMWDDTAT-TLTATNIVG 509  
Db 728 WLKGRPLSLTE--RHHFTP-CNQLLVQNV-----MIDDAGRYTCENSNPLG 772  
Qy 510 CAQFFTKADY---POCQSMHSQILGTMILVIGIIVATLLVFIIVILMVRYKVCNHEAPS 566  
Db 773 TERAHSQLSILPTPGCRK-DGTTVGIFTIIVVCSIVLTSLVWVCIIYQTRKK----- 823  
Qy 567 KWAAAVSNVYSQTNGAQPSPSSAGAPQGP-----PKVVVYRNE----- 607  
Db 824 -----SEESYVTNTDETITVPPDVPVSYLSSQGTLSDRQETVVRTEGHOANGHIESNGVC 877  
Qy 608 LLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPR-PKPSLDRLMGAF-----AS 662  
Db 878 LRD--PSLFEYVDIHSITTCRQPKLCVGYTRPQWYTERKADRTAAPHTTAHSGSAVSCDCS 935  
Qy 663 LDLKSRKEELDSRTTPAGRGAGTSARGHSHSDREPLGPPAARASRLPLPLEGKAKRSH 722  
Db 936 TDTAYHPQVPRDSGQP-GTASSQELRQHDREYSP-----HHPSYSGTADGSH 981  
Qy 723 SEDMGDFAAAAAGGVVPGYSPRKVSNLWTKESLVNGMLLPFEESDLVGARTGTFGSSE 782  
Db 982 TL-----SGGSLYPSNHD-----RIPLSLKNKAASADGN-GDSS 1014  
Qy 783 WVM 785  
Db 1015 WTL 1017

RESULT 3

US-08-986-485-2  
; Sequence 2, Application US/08986485  
; Patent No. 6046030  
; GENERAL INFORMATION:  
; APPLICANT: WU, SHUJIAN  
; APPLICANT: SWEET, RAYMOND  
; APPLICANT: TRUNER, ALEMGESD  
; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RATNER & PRESTIA  
; STREET: P.O. BOX 980  
; CITY: VALLEY FORGE  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19482  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/986,485  
; FILING DATE: 08-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/059,448  
; FILING DATE: 22-SEP-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PRESTIA, PAUL F  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GH-70264  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1101 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-986-485-2

Query Match 8.6%; Score 351; DB 3; Length 1101;  
Best Local Similarity 21.6%; Pred. No. 1.8e-20;  
Matches 168; Conservative 77; Mismatches 233; Indels 298; Gaps 23;

QY 30 LSESLGTLCPKSGLL-----FVPPDIDRRVTVELRLGNGFIHISRQDFANMTGLVDLT 82  
DB 193 LRSLLTLRLSKNRITQLPVRAFKLP-----RLTQDLNLRNRLIELGLTFQGLNSLEVLUK 248  
QY 83 LSRNTISHTQPSFPLESLRSLHLSNRLPSLGEDTLRLGLVNLQHLIVNNQLGGLADE 142  
DB 249 LQNNISKLTDGAFWGLSKMHLHLEYDSLVEVNSGLYGLTALHQLHLSNNSIARIHRK 308  
QY 143 AFEDFLLTLEDLDSLYNNHGLPWDSVRRMVNLHQLSLDHNLDDHTAEGTFADLQKLARL 202  
DB 309 GW-SFCQKLHGLVLSFNNITLDEESLAEVLSVLRSLSHNSISHTAEGAFKGLSLRVL 367  
QY 203 DLTSNRLQKLPPDP--IFARSOASALTATPFAPPLS-----FSFGN--- 242  
DB 368 LDHNEISGTIEDTSCAFSGLEFGCHSKLTLFGNKIKSVAKRAPSGLEGLEHLNLGNATR 427  
QY 243 -----PLH-----CNCELLWL-----RRLERDDDLCTCGSPGLKGRYF 276  
DB 428 SVQDFAFVKMKNLKELHISSDSFLCDQLKWLPPWLIQRLQAFVTTATCAHPESLKGQSI 487  
QY 277 WHVREEEFVCE-----PPLITQ----- 293  
DB 488 FSVPPSFVCDLFLKPOIITQETTMMVGMKDIRFTCSAASSSSSPMTFAWKDNNEVLTN 547

QY 294 -----HTHK----- 297  
DB 548 ADMENFVHVAQDGEVMEYTTILHLRQVTFGHEGRYQCVITNHFSGTYSKARLTNVNLP 607  
QY 298 -----LLVLEGQAATLKCKAIGDPSPLHW-----VAPDDDELGVN 332  
DB 608 SFTKTPHDITRTTTVARLECAATGHPNPQIAWKDGGTDFPAARERRMHVMPDD----- 662  
QY 333 SSRTAVYDNGTLDIFITTS---QDSGAFCTIAANAAGATAMVEVSIIVOLPHLSNSTRT 389  
DB 663 -----DVFFITDKIDDDAGVYSCTAQNSAGSISANATLTVLETPSL-----V 704  
QY 390 APPKSRSLSDITGSSKTSRGGSGGGEPPKSPERAVLVSEVTTTSALVWKSVSAPRV 449  
DB 705 VPLEDR-----VVSVGETVALOCKATGNPPRI 732  
QY 450 KMYQLQYNCSDDEVILYRMIPASNKAFVNNLVSGTGYDLCVLAMWDDDTATTLTATNIVG 509  
DB 733 TWFKGDRPLSLTE--RHHLTP--DNQLLVVQNVVAE-----DAGRYTCMSNTLIG 778  
QY 510 CAQFFTKADYPQCQSMHSQI-----LGTMILVIGGIIIVATLLVFIILMV 555  
DB 779 -----TERAHSQLSVLPAAGCKKGGTGTGIFTIAVSSIVLTSLVWVCIIYQT 826  
QY 556 RYKVCNHEAPSXMAAAVSNVYSQTNQAQPPPPSSAPAGAPQGP-----PKVVVRNE 607  
DB 827 RKK-----SEESVTTNTDETVPDPVPSVLSQGTLSRQETVURTE 868

RESULT 4  
US-09-063-950-2  
; Sequence 2, Application US/09063950C  
; Patent No. 6225085  
; GENERAL INFORMATION:  
; APPLICANT: Holtzman, Douglas A.  
; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE OF INVENTION: THEREFOR  
; FILE REFERENCE: MEI-019  
; CURRENT APPLICATION NUMBER: US/09/063,950C  
; CURRENT FILING DATE: 1998-04-21  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 673  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-063-950-2

Query Match 7.3%; Score 299; DB 4; Length 673;  
Best Local Similarity 23.7%; Pred. No. 1.8e-16;  
Matches 169; Conservative 77; Mismatches 268; Indels 198; Gaps 28;

QY 8 LLAFGMAFVVDACPKYCVQCNLSLGLTCLPCKSGLLFVPPDIDRRVTVELRLGNGFIH 67  
DB 14 LLALGPG---VQCGPSGQC---SQPTVCTARQTTVPDPVPPDPTVGLYFENGITML 67  
QY 68 SQDFANMTGLVDLTLSRNTISHIQPFSLDLESRLSLHLSNRLPSLGEDTLRLGLVNLQ 127  
DB 68 DAGSFAGUGLQGLDLSQNIASLPSGVFQPLANISNLDLTANRLHEITNETFRGLRUE 127  
QY 128 HLIVNNQLGGTADBAFEDFLTLEDLDSLYNNHGLPWDSVRRMVNLHQLSLDHNL--- 184  
DB 128 RLYLCKNRIRHTOPGAF-DTLRLLELKLQDLNELRALP---PLRPRLLLLDLSHNSLLA 183  
QY 185 -----LDHIAEGTFADLQKLARLDTNRLQKLPPD----- 215  
DB 184 LEPGILDTANVEALRAGLGLQQLDEGLFSRLNRLHDLVDSDNQLERVPVIRGLRGLTR 243  
QY 216 -----PIFARSOASALTATPFAPPLSFSF-----GGNPLHCN 247  
DB 244 LRLAGNTRIAQURPEDLAGLAALQELVDVSNLSLQALPGDLGLFPRLRLAAARNPFCV 303

QY 248 CELLMLRLERDDDDLETCGSPGGLKGRYFHWVREBEFVC 306  
Db 304 CPLSWFGPWVRESHV-TLASP-----BETRCHFP-----KNAGRL--LELYA 345  
QY 307 TLKCKA-----IGDPSPLIH-----WVAPDRLVGSSRTAVDNGTLDFI 348  
Db 346 DFGCAPTTTATVPTTRPVVREPTALSSLATWLS----- 382  
QY 349 TTSQDSGAFCTIAANA-AGEATAMVEVSIQVPHLSNSTSTAPPKSRIS----- 397  
Db 383 -----TAPATEAPSPSTAPPTVGPVQP-----QDCPPSTCLNGTCHLGRH 426  
QY 398 -----DITGSSKTSRGSGSGGEPKSP-PERAVL--VSEVTTTSALV--KWSYSK 444  
Db 427 HLACLCPEFTGLYCESONGQSTPPTVTRPRPSRLTGLIEPVSPSLRVLQRYLQ 486  
QY 445 SAPRVKMYOLYN--CSDDEVLIYRIMPASNAFVNNILVSGTGYDLCLVAMDDTATTL 502  
Db 487 SSVQLSRILTYRNLSPDKRLVTLRLPASLAETVTLRPNATYSCVCMPL---GPG 542  
QY 503 TATNIVGCAOFFTKADYPOCOSMH---SOILGTMILVIGGIIVATLLVFIILMVRYKV 559  
Db 543 VPGEAEACGEAHTP---PAVHSNHAPVTOAREGNPLLIAPALAAVLLAALAAVGAAYCV 599  
QY 560 CNHEAFSKMAAASVNVYSQTNGAOP-----PPSSAPAG--APPOG 598  
Db 600 RRGRA---MAAAODKQGVGPGAGLEBKVLPBPGPKATEGGEALPSG 648

RESULT 5  
US-09-191-647-2  
; Sequence 2, Application US/09191647  
; Patent No. 6046015  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey  
; APPLICANT: Kid, Thomas  
; APPLICANT: Brose, Katja  
; APPLICANT: Tessier-Lavigne, Marc  
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
; FILE REFERENCE: B98-031-3  
; CURRENT APPLICATION NUMBER: US/09/191,647  
; CURRENT FILING DATE: 1998-11-13  
; EARLIER APPLICATION NUMBER: 60/065,544  
; EARLIER FILING DATE: 1997-11-14  
; EARLIER APPLICATION NUMBER: 60/081,057  
; EARLIER FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1525  
; TYPE: PRT  
; ORGANISM: human  
US-09-191-647-2  
Query Match 6.9%; Score 284.5; DB 3; Length 1525;  
Best Local Similarity 29.2%; Pred. No. 1.1e-14;  
Matches 84; Conservative 36; Mismatches 117; Indels 51; Gaps 6;  
QY 9 LAFGMAFAVD-----ACPKYCVCONLSGTLCPKSGLLFVPPDIDRRTVELRLGNF 63  
Db 11 LSLGLVLAILNKVAPQACQSCSGSTVD---CHGLALRSVPRNIPRNTERLDLNGN 66  
QY 64 IIHISROPANMTGLVDLTLSRNTTISHIOPFSLDLSRLHLSNRLPSLGEDTLRGL 123  
Db 67 ITRITKTDFAGLHRLVQLMENKISTIERGAFQDLKELERLNRNHLQFPPELLFLT 126  
QY 124 VNLQHLIVNNQLGGIADEAFEDFLTLEDLSYNNHLGLPWSVRRMNLHOLSLDHN 183  
Db 127 AKLYR-----LDLSENQIOAIPKRAFGAVDIKNQLDYN 161  
QY 184 LLDHIAEGTFADLOKLARLDLTNSRLQKLPDPPIFARSOASALTATPPAPPL-SFSFGN 242  
Db 162 QISCIEDGAFRALDLEVLTLNNNITRL-----SVASFNMHMKLRTFRLHSN 209

QY 243 PLHCNCELL-----WLRRLERDDDDLETCGSPGGLKGRYFHWVREBEFVC 286  
Db 210 NLYCDCHLAWLSDLWLRKRPVGLYTQCMGSPHLRGHNVAEVQKREFVC 257  
RESULT 6  
US-09-540-245A-2  
; Sequence 2, Application US/09540245A  
; Patent No. 6270984  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey  
; APPLICANT: Kid, Thomas  
; APPLICANT: Brose, Katja  
; APPLICANT: Tessier-Lavigne, Marc  
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
; FILE REFERENCE: B98-031-3  
; CURRENT APPLICATION NUMBER: US/09/540,245A  
; CURRENT FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 60/065,544  
; PRIOR FILING DATE: 1997-11-14  
; PRIOR APPLICATION NUMBER: 60/081,057  
; PRIOR FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1525  
; TYPE: PRT  
; ORGANISM: human  
US-09-540-245A-2  
Query Match 6.9%; Score 284.5; DB 4; Length 1525;  
Best Local Similarity 29.2%; Pred. No. 1.1e-14;  
Matches 84; Conservative 36; Mismatches 117; Indels 51; Gaps 6;  
QY 9 LAFGMAFAVD-----ACPKYCVCONLSGTLCPKSGLLFVPPDIDRRTVELRLGNF 63  
Db 11 LSLGLVLAILNKVAPQACQSCSGSTVD---CHGLALRSVPRNIPRNTERLDLNGN 66  
QY 64 IIHISROPANMTGLVDLTLSRNTTISHIOPFSLDLSRLHLSNRLPSLGEDTLRGL 123  
Db 67 ITRITKTDFAGLHRLVQLMENKISTIERGAFQDLKELERLNRNHLQFPPELLFLT 126  
QY 124 VNLQHLIVNNQLGGIADEAFEDFLTLEDLSYNNHLGLPWSVRRMNLHOLSLDHN 183  
Db 127 AKLYR-----LDLSENQIOAIPKRAFGAVDIKNQLDYN 161  
QY 184 LLDHIAEGTFADLOKLARLDLTNSRLQKLPDPPIFARSOASALTATPPAPPL-SFSFGN 242  
Db 162 QISCIEDGAFRALDLEVLTLNNNITRL-----SVASFNMHMKLRTFRLHSN 209

QY 243 PLHCNCELL-----WLRRLERDDDDLETCGSPGGLKGRYFHWVREBEFVC 286  
Db 210 NLYCDCHLAWLSDLWLRKRPVGLYTQCMGSPHLRGHNVAEVQKREFVC 257  
RESULT 7  
US-09-540-153-2  
; Sequence 2, Application US/09540153  
; Patent No. 6270995  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey  
; APPLICANT: Kid, Thomas  
; APPLICANT: Brose, Katja  
; APPLICANT: Tessier-Lavigne, Marc  
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
; FILE REFERENCE: B98-031-3  
; CURRENT APPLICATION NUMBER: US/09/540,153  
; CURRENT FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 60/065,544  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: 60/081,057  
; PRIOR FILING DATE: 1998-04-07

```
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: human
US-09-540-153-2

Query Match
Best Local Similarity 6.9%; Score 284.5; DB 4; Length 1525;
Matches 84; Conservative 36; Mismatches 117; Indels 51; Gaps 6;

QY 9 LAFGMAFAVVD-----ACPKYCVQNLSLGLTLCPSKGLLFVPPDIDRRTVELRLGNGF 63
DB 11 LSLGLVLAITLNKVAQACQAQSCSGSTVD-----CHGLALRSVPRNIPNTERLDLNGNN 66
QY 64 IHHISRODFANMTGLVDLTLSRNTTISHIQPFSLDLSRLSLHLSNRLPSLGEDTLRGL 123
DB 67 ITRITKDFAGRLHRLVQLMENKISTIERGAFODLKELERLRLNRNHLQLPPELLFLGT 126
QY 124 VNLQHLIVNNQLGGIADAEFEDFLTLEDLDSYNNHLGHPWDSVRRMVLHQLSLDHN 183
DB 127 AKLYR-----LDLSENQIAIPKAPRGAVIDIKNLQLDYN 161
QY 184 LLDHIAEGTFADLOKLARLDLTLSNRLOKLPDPPIFARSOASALTATPAPPL-SFSFGN 242
DB 162 QTSICIEDGAFRALDRDLVLTNNNITRL-----SVASFNMHPKLTTPRLHSN 209
QY 243 PLHCNCELL-----WLRLERDDLETGSGPGLKGRYFWHVRREEFVC 286
DB 210 NLYCDCHLAWLSDWLKRPVGLYTQCMGPHLRGHNAEVOVKREFVC 257

RESULT 8
US-09-182-024A-2
; Sequence 2, Application US/09182024A
; Patent No. 6342370
; GENERAL INFORMATION:
; APPLICANT: Connolly, Timothy
; APPLICANT: Rajput, Bhanu
; TITLE OF INVENTION: Human Slit Polypeptide and Polynucleotides Encoding
; TITLE OF INVENTION: Same
; FILE REFERENCE: 640100-271
; CURRENT APPLICATION NUMBER: US/09/182,024A
; CURRENT FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/063,946
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/096,420
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1523
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-182-024A-2

Query Match
Best Local Similarity 6.5%; Score 268; DB 4; Length 1523;
Matches 86; Conservative 41; Mismatches 130; Indels 64; Gaps 7;

QY 9 LAFGMAFAV-----DAPKYCVQNLSLGLTLCPSKGLLFVPPDIDRRTVELRLGNG 62
DB 16 LALALALASVLGPPAVACPTKCTCSAASVD-----CHGLGLRAVPRGIPRANERLDLDRN 71
QY 63 FHHISRODFANMTGLVDLTLSRNTTISHIQPFSLDLSRLSLHLSNRLPSLGEDTLRGL 122
DB 72 NTRITKDFAGRLHRLVQLMENKISTIERGAFODLKELERLRLNRNHLQLPPELLFOS 131
QY 123 LVNLQHLIVNNQLGGIADAEFEDFLTLEDLDSYNNHLGHPWDSVRRMVLHQLSLDHN 182
DB 132 TPFLTR-----LDLSENQIQIPKAPRGAVIDIKNLQLDYN 166

; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: human
US-09-540-153-2

Query Match
Best Local Similarity 6.9%; Score 284.5; DB 4; Length 1525;
Matches 84; Conservative 36; Mismatches 117; Indels 51; Gaps 6;

QY 9 LAFGMAFAVVD-----ACPKYCVQNLSLGLTLCPSKGLLFVPPDIDRRTVELRLGNGF 63
DB 11 LSLGLVLAITLNKVAQACQAQSCSGSTVD-----CHGLALRSVPRNIPNTERLDLNGNN 66
QY 64 IHHISRODFANMTGLVDLTLSRNTTISHIQPFSLDLSRLSLHLSNRLPSLGEDTLRGL 123
DB 67 ITRITKDFAGRLHRLVQLMENKISTIERGAFODLKELERLRLNRNHLQLPPELLFLGT 126
QY 124 VNLQHLIVNNQLGGIADAEFEDFLTLEDLDSYNNHLGHPWDSVRRMVLHQLSLDHN 183
DB 127 AKLYR-----LDLSENQIAIPKAPRGAVIDIKNLQLDYN 161
QY 184 LLDHIAEGTFADLOKLARLDLTLSNRLOKLPDPPIFARSOASALTATPAPPL-SFSFGN 242
DB 162 QTSICIEDGAFRALDRDLVLTNNNITRL-----SVASFNMHPKLTTPRLHSN 209
QY 243 PLHCNCELL-----WLRLERDDLETGSGPGLKGRYFWHVRREEFVC 286
DB 210 NLYCDCHLAWLSDWLKRPVGLYTQCMGPHLRGHNAEVOVKREFVC 257

RESULT 9
US-09-191-647-7
; Sequence 7, Application US/09191647
; Patent No. 6046015
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brosse, Katja
; APPLICANT: Tessier-Lavigne, Marc
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/191,647
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: 60/065,544
; EARLIER FILING DATE: 1997-11-14
; EARLIER APPLICATION NUMBER: 60/081,057
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1480
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-191-647-7

Query Match
Best Local Similarity 6.4%; Score 263; DB 3; Length 1480;
Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;

QY 6 GGLLAFGMFAF-----AVVDA-CPKYCVQNLSLGLTLCPSKGLLFVPPDIDRRTVEUR 58
DB 51 GGLGSVGHIHPGGVGIVTEARCPVCSCTGLNVD-----CSHRLTSPVKISADVERLE 106
QY 59 LGGNFIIHISRODFANMTGLVDLTLSRNTTISHIQPFSLDLSRLSLHLSNRLPSLGED 118
DB 107 LQNNLTVIYETDFORLTKLRLQTDNQIHTIERNSFODLVSLERLDSNNVITTVGRR 166
QY 119 TLRLGLVNLQHLIVNNQLGGIADAEFEDFLTLEDLDSYNNHLGHPWDSVRRMVLHQL 178
DB 167 VFKGAQSLSRLQDNNQITCLDEHAFKG-LVELEILTNNNNLTSLP-----212
QY 179 SLDNHLDHIAEGTFADLOKLARLDLTLSNRLOKLPDPPIFARSOASALTATPAPPLSFS 238
DB 213 ---HNI-----FGLGLRALRLSDN-----BFA-----233
QY 239 FGGNPLHNCNCELLWLRRLERD-----DDLETGSGPGLKGRYFWHVRREEFVCEPPLITQH 294
DB 234 -----CDCHLSLWSLRSFLRSATRLAPYTRCOSPSQLKQNVADLHDQBFKCSG--LTEH 284
QY 295 THKLLVLEGQAATLKCKA 312
DB 285 -----APMECGA 291

RESULT 10
US-09-540-245A-7
; Sequence 7, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
```

APPLICANT: Kid, Thomas  
APPLICANT: Brose, Katja  
APPLICANT: Tessier-Lavigne, Marc  
TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
FILE REFERENCE: B98-031-3  
CURRENT APPLICATION NUMBER: US/09/540,245A  
CURRENT FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 60/065,544  
PRIOR FILING DATE: 1997-11-14  
PRIOR APPLICATION NUMBER: 60/081,057  
PRIOR FILING DATE: 1998-04-07  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 1480  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-540-245A-7

Query Match 6.4%; Score 263; DB 4; Length 1480;  
Best Local Similarity 28.0%; Pred. No. 6.3e-13;  
Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;  
QY 6 GGLAFGMAF-----AVVDA-CPKYCVCONLSGLTCLPSKGLLFVPPDIDRRITVELR 58  
DB 51 GGLSGVGIHPGGVGVIETEARCPVCSCTGLNVD-----CSHRLTSVPRKISADVERLE 106  
QY 59 LGGNFIHISRODFANMTGLVDLTLSRNTISHIQPFSFLDLSRLSLHLDNSRLPSLGED 118  
DB 107 LQGNLTVIYETDFQRLTKRLMLQLTNDQIHTIERNFSQDLVSLERLDSINNVTITVGR 166  
QY 119 TLRLGLVNLQHLIVNNNOLGGIADAEFLLTLEDLDSYNNLHGLPWSVRRMVNLHQL 178  
DB 167 VFKAQSLRSLOLDNNQITCLDEHAFKG-LVELEILTNNNLTSLP----- 212  
QY 179 SLDNHLLDHIAGCTFADLQKLARLDLTNSRLQKLPDPPIFARSOASALTATPAPPLSFS 238  
DB 213 ---HNI-----FGGLRLRALRLSDN-----PFA----- 233  
QY 239 FGGNPLHCNCELLWLRLERD-----DLETGSPGGLKGRYFHWVREBEFVCEPPLITQH 294  
DB 234 -----CDCHLSWLSRFLRSATRLAPYTRCQSPQLKQNVADLHDOEFKCSG--LTEH 284  
QY 295 THKLLVLEGOAATLKCKA 312  
DB 285 -----APMECGA 291

RESULT 11  
US-09-540-153-7  
Sequence 7, Application US/09540153  
Patent No. 6270995  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey  
APPLICANT: Kid, Thomas  
APPLICANT: Brose, Katja  
APPLICANT: Tessier-Lavigne, Marc  
TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
FILE REFERENCE: B98-031-3  
CURRENT APPLICATION NUMBER: US/09/540,153  
CURRENT FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 09/191,647  
PRIOR FILING DATE: 1998-11-13  
PRIOR APPLICATION NUMBER: 60/081,057  
PRIOR FILING DATE: 1998-04-07  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 1480  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-540-153-7

Query Match 6.4%; Score 263; DB 4; Length 1480;  
Best Local Similarity 28.0%; Pred. No. 6.3e-13;  
Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;  
QY 6 GGLAFGMAF-----AVVDA-CPKYCVCONLSGLTCLPSKGLLFVPPDIDRRITVELR 58  
DB 51 GGLSGVGIHPGGVGVIETEARCPVCSCTGLNVD-----CSHRLTSVPRKISADVERLE 106  
QY 59 LGGNFIHISRODFANMTGLVDLTLSRNTISHIQPFSFLDLSRLSLHLDNSRLPSLGED 118  
DB 107 LQGNLTVIYETDFQRLTKRLMLQLTNDQIHTIERNFSQDLVSLERLDSINNVTITVGR 166  
QY 119 TLRLGLVNLQHLIVNNNOLGGIADAEFLLTLEDLDSYNNLHGLPWSVRRMVNLHQL 178  
DB 167 VFKAQSLRSLOLDNNQITCLDEHAFKG-LVELEILTNNNLTSLP----- 212  
QY 179 SLDNHLLDHIAGCTFADLQKLARLDLTNSRLQKLPDPPIFARSOASALTATPAPPLSFS 238  
DB 213 ---HNI-----FGGLRLRALRLSDN-----PFA----- 233  
QY 239 FGGNPLHCNCELLWLRLERD-----DLETGSPGGLKGRYFHWVREBEFVCEPPLITQH 294  
DB 234 -----CDCHLSWLSRFLRSATRLAPYTRCQSPQLKQNVADLHDOEFKCSG--LTEH 284  
QY 295 THKLLVLEGOAATLKCKA 312  
DB 285 -----APMECGA 291

RESULT 12  
US-09-182-024A-5  
Sequence 5, Application US/09182024A  
Patent No. 6342370  
GENERAL INFORMATION:  
APPLICANT: Connolly, Timothy  
APPLICANT: Rajput, Bhanu  
TITLE OF INVENTION: Human Slit Polypeptide and Polynucleotides Encoding  
TITLE OF INVENTION: Same  
FILE REFERENCE: 640100-271  
CURRENT APPLICATION NUMBER: US/09/182,024A  
CURRENT FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: 60/063,946  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/096,420  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 1480  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-182-024A-5

Query Match 6.4%; Score 263; DB 4; Length 1480;  
Best Local Similarity 28.0%; Pred. No. 6.3e-13;  
Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;  
QY 6 GGLAFGMAF-----AVVDA-CPKYCVCONLSGLTCLPSKGLLFVPPDIDRRITVELR 58  
DB 51 GGLSGVGIHPGGVGVIETEARCPVCSCTGLNVD-----CSHRLTSVPRKISADVERLE 106  
QY 59 LGGNFIHISRODFANMTGLVDLTLSRNTISHIQPFSFLDLSRLSLHLDNSRLPSLGED 118  
DB 107 LQGNLTVIYETDFQRLTKRLMLQLTNDQIHTIERNFSQDLVSLERLDSINNVTITVGR 166  
QY 119 TLRLGLVNLQHLIVNNNOLGGIADAEFLLTLEDLDSYNNLHGLPWSVRRMVNLHQL 178  
DB 167 VFKAQSLRSLOLDNNQITCLDEHAFKG-LVELEILTNNNLTSLP----- 212  
QY 179 SLDNHLLDHIAGCTFADLQKLARLDLTNSRLQKLPDPPIFARSOASALTATPAPPLSFS 238  
DB 213 ---HNI-----FGGLRLRALRLSDN-----PFA----- 233



Query Match	6.4%	Score 263;	DB 5;	Length 1480;
Best Local Similarity	28.0%	Pred. No. 6.3e-13;		
Matches	89;	Conservative	39;	Mismatches 102;
				Indels 88;
				Gaps 11;

```
Query Match      6.4%; Score 263; DB 5; Length 1480;
Best Local Similarity 28.0%; Pred. No. 6.3e-13;
Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;
```

---

Qy	6	GGLAFGMAF-----AVDA-CPKYVCVQNLSLGTCLCPKSGLLFPVPDDIRRTVELR	58
Dd	51	GGLGSVGIHPTGGVGVIPEARCVCSCTGLNVD---CXHRGLTSVPFKISADVERLE	106
Qy	59	LGNFNFIHSRODFANMTGLVDUTLGRNTISHIQPFSLDLESRLSHLDNSLRPLSGLD	118
Dd	107	LOGNNLTIVETDFQRILTKRLMLQLTDNQIHTIERNSFDLVSLERLIDISNNVITTVRR	166
Qy	119	TLRGVLNLQHILIYNVNOLGCIADAEDFLTLIEDLDLSYNNHLGLPWDSVRRMVNLHQ	178
Dd	167	VFKGAOSRLSDLNQNIQTCDHEAPKG-LVEILEILTNNNITSLP-----	212
Qy	179	SLDHNLLDHTAEGTFADLOKLARLDLITSNRLOKLPPDPIFARSQASALTATPEAPPLSF	238
Dd	213	--HNI-----FGGLGRLARLRUSDN-----PFA-----	233
Qy	239	FGGNPLHCNCCELLWLRRLRD----DDLETGCGPGGLKGRYPWHVREEEEVCPPPLITOH	294
Dd	234	-----CDCHLSWSFFLRSATRELAPYTQCSPSQLKGONVADLHDQEFKCQG--LTEH	284
Qy	295	THKLLVLEGQAATLKCKKA 312	
Dd	285	-----APMECCA 291	

RESULT 14  
PCT-US91-09055-3  
; Sequence 3, Application PC/TUS9109055  
; GENERAL INFORMATION:  
; APPLICANT: Rothberg, Jonathan Marc and Artavanis-Taikonas, Spyridon  
; TITLE OF INVENTION: Purified SLIT protein and Sequence Elements Thereof  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Yale University  
; ADDRESSEE: Office of Cooperative Research  
; STREET: 246 Church Street  
; STREET: Suite 401  
; CITY: New Haven  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 06510  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch. 800 Kb storage  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh 6.0.5  
; SOFTWARE: Microsoft Word 4.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/09055  
; FILING DATE: 19911127  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/624,135  
; FILING DATE: 7-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barth, Richard J.  
; REGISTRATION NUMBER: 28,180  
; REFERENCE/DOCKET NUMBER: 900964/RSB  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 972-1400  
; TELEFAX: (212) 370-1622  
; TELEX: 236268  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 222 amino acids  
; TYPE: AMINO ACIDS

LENGTH: 231 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-986-485-7

Query Match  
Best Local Similarity 29.1%; Score 259; DB 3; Length 231;  
Matches 81; Conservative 33; Mismatches 94; Indels 70; Gaps 8;

QY 21 CPKYCVCONSESLGTLCPKGLFVPPDIDRRTVELRLGNGFIHISRQDFANWTGLVD 80  
DB 17 CPRVCSTGLNVD---CSHRGLTSVPRKISADVERLELOGNLTVIYETDFQRLTKLRM 72

QY 81 LTLRSNTISHIQPFSELDLSRLHLDNSRLPSLGEDTLRGLVNLQHLIVNNNOLGGIA 140  
DB 73 LQTDNQIHTIERNISFQDLVSLERLDSINNVTIVGRVFKGAQSLRSLQDNNQITCLD 132

QY 141 DEAFEDFLITLEDLDLSYNNLHGLPWSVRRMVNLHQLSLDNLHDHIAEGTFADLQKLA 200  
DB 133 EHAFKG-LVELEILTNNNNLTSLP-----HNI-----FGGLGLRLR 167

QY 201 RDLTSLNRLQKLPDPPIFARSOASALTATPEAPPLSFSGGNPLHCNCELLWRLRLRD- 259  
DB 168 ALRLSDN-----PFA-----CDCHLSWLSRFLRSA 192

QY 260 ---DDLETGSPGGLKGRYFWHVREEFVCEPPLITQH 294  
DB 193 TRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 228

Search completed: June 8, 2003, 21:32:09  
Job time : 41 secs

TOPOLOGY: Linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Plank-LRR-Plank 1  
LOCATION: 1 to 222  
IDENTIFICATION METHOD: similarity to other Plank-LRR-  
OTHER INFORMATION: mediates adhesive events  
PCT-US91-09055-3

Query Match  
Best Local Similarity 28.4%; Score 260.5; DB 5; Length 222;  
Matches 84; Conservative 35; Mismatches 96; Indels 81; Gaps 9;

QY 21 CPKYCVCONSESLGTLCPKGLFVPPDIDRRTVELRLGNGFIHISRQDFANWTGLVD 80  
DB 1 CPRVCSTGLNVD---CSHRGLTSVPRKISADVERLELOGNLTVIYETDFQRLTKLRM 56

QY 81 LTLRSNTISHIQPFSELDLSRLHLDNSRLPSLGEDTLRGLVNLQHLIVNNNOLGGIA 140  
DB 57 LQTDNQIHTIERNISFQDLVSLERLDSINNVTIVGRVFKGAQSLRSLQDNNQITCLD 116

QY 141 DEAFEDFLITLEDLDLSYNNLHGLPWSVRRMVNLHQLSLDNLHDHIAEGTFADLQKLA 200  
DB 117 EHAFKG-LVELEILTNNNNLTSLP-----HNI-----FGGLGLRLR 151

QY 201 RDLTSLNRLQKLPDPPIFARSOASALTATPEAPPLSFSGGNPLHCNCELLWRLRLRD- 259  
DB 152 ALRLSDN-----PFA-----CDCHLSWLSRFLRSA 176

QY 260 ---DDLETGSPGGLKGRYFWHVREEFVCEPPLITQHTHKLIVLEGQAATLKCKA 312  
DB 177 TRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH-----APMECGA 219

RESULT 15  
US-08-986-485-7  
Sequence 7, Application US/08986485  
Patent No. 6046030  
GENERAL INFORMATION:  
APPLICANT: WU, SHUJIAN  
APPLICANT: SWEET, RAYMOND  
APPLICANT: TRUNEH, ALEMESEGED  
TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: RATNER & PRESTIA  
STREET: P.O. BOX 980  
CITY: VALLEY FORGE  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/986.485  
FILING DATE: 08-DEC-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/059,448  
FILING DATE: 22-SEP-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: PRESTIA, PAUL F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 21:30:28 ; Search time 55 Seconds  
(without alignments)  
1481.029 Million cell updates/sec

Title: US-09-831-846-2  
Perfect score: 4094  
Sequence: 1 METLLGLLAFGMFAVVD.....DLVGARGTGSSEWYMSTV 789

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW PUB.pap.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW PUB.pap.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW PUB.pap.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW PUB.pap.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW PUB.pap.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW PUB.pap.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pap.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW PUB.pap.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1566	38.3	628	10	US-09-815-626-2
2	1563.5	38.2	627	9	US-10-245-103-100
3	1563.5	38.2	627	9	US-10-245-107-100
4	1563.5	38.2	627	9	US-10-245-143-100
5	1563.5	38.2	627	9	US-10-245-771-100
6	1563.5	38.2	627	9	US-10-245-851-100
7	1563.5	38.2	627	9	US-10-245-883-100
8	1563.5	38.2	627	9	US-10-237-535-100
9	1563.5	38.2	627	9	US-10-238-183-100
10	1563.5	38.2	627	9	US-10-238-283-100
11	1563.5	38.2	627	9	US-10-238-370-100
12	1563.5	38.2	627	9	US-10-245-055-100
13	1563.5	38.2	627	9	US-10-245-147-100
14	1563.5	38.2	627	9	US-10-245-730-100
15	1563.5	38.2	627	9	US-10-245-739-100
16	1563.5	38.2	627	9	US-10-246-210-100
17	1563.5	38.2	627	9	US-10-239-196-100
18	1563.5	38.2	627	9	US-10-243-024-100
19	1563.5	38.2	627	9	US-10-243-409-100

20	1563.5	38.2	627	9	US-10-245-033-100	Sequence 100, App
21	1563.5	38.2	627	9	US-10-245-621-100	Sequence 100, App
22	1563.5	38.2	627	9	US-10-245-880-100	Sequence 100, App
23	1563.5	38.2	627	9	US-10-243-095-100	Sequence 100, App
24	1563.5	38.2	627	9	US-10-245-185-100	Sequence 100, App
25	1563.5	38.2	627	9	US-10-245-427-100	Sequence 100, App
26	1563.5	38.2	627	9	US-10-245-473-100	Sequence 100, App
27	1563.5	38.2	627	9	US-10-245-770-100	Sequence 100, App
28	1563.5	38.2	627	9	US-10-245-877-100	Sequence 100, App
29	1563.5	38.2	627	9	US-10-246-976-100	Sequence 100, App
30	1563.5	38.2	627	9	US-10-243-320-100	Sequence 100, App
31	1563.5	38.2	627	9	US-10-243-743-100	Sequence 100, App
32	1563.5	38.2	627	9	US-10-242-845-100	Sequence 100, App
33	1545.5	37.8	551	10	US-09-897-214-8	Sequence 8, Appli
34	1454	35.5	565	10	US-09-768-826-58	Sequence 58, Appl
35	1144	27.9	468	10	US-09-768-826-40	Sequence 40, Appl
36	452	11.0	713	9	US-10-028-392-35	Sequence 35, Appl
37	451	11.0	713	10	US-09-822-687-2	Sequence 2, Appli
38	441	10.8	634	9	US-10-028-392-2	Sequence 2, Appli
39	435.5	10.6	640	9	US-09-905-291A-292	Sequence 292, App
40	435.5	10.6	640	9	US-09-992-598-501	Sequence 501, App
41	435.5	10.6	640	9	US-09-989-293A-501	Sequence 501, App
42	435.5	10.6	640	9	US-09-902-853-292	Sequence 292, App
43	435.5	10.6	640	9	US-09-989-735-501	Sequence 501, App
44	435.5	10.6	640	9	US-09-990-444-501	Sequence 501, App
45	435.5	10.6	640	9	US-10-114-893-117	Sequence 117, App

ALIGNMENTS

RESULT 1  
US-09-815-626-2  
; Sequence 2, Application US/09815626  
; Patent No. US20020076752A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria Alexandra  
; TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT  
; TITLE OF INVENTION: FAMILY MEMBER AND USES-THEREOF  
; FILE REFERENCE: 10448-031001  
; CURRENT APPLICATION NUMBER: US/09/815,626  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/191,863  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-815-626-2

Query Match	38.3%	Score	1566	DB	10	Length	628
Best Local Similarity	52.9%	Pred. No.	1.2e-78				
Matches	311	Conservative	88	Mismatches	175	Indels	14
Gaps	5						
Qy	16	AVVDACPKVCVCONLSESLGTLCPKSGLLFPVPPDIDRTVELRLGNGFIHISRODFANM	75				
Db	23	ATPSPCPRRCRCQTQSLFSLVLCPCAGLLFPVPSLDRAAEALRLADNFIAVRRDLANM	82				
Qy	76	TGLVDTLSRNTISHIQPFSLDLSRLSLHLDNSRLPSLGBEDTLRGLVNLQHLIVNNQ	135				
Db	83	TGLHLLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSNNQ	142				
Qy	136	LGGIADAFAEDPFLTLLELDLSYNNLHGLPWSVVRMWNHOLSLDHLLDHIAGTFAD	195				
Db	143	LAALAAGALDDCAETLELDLSYNNLEQLPWEALGLGNVNTLGLDHLNLSASVPAGAFSR	202				
Qy	196	LQKARLDLTNSRLQKLPDPFARSOASALTATPAPPLS---FSFGGNPLHNCCELLW	252				
Db	203	LHKARLDMTNSRLTIIPDPPLFSR---LPLLARPGSPASALVIAFGNPLHNCCELLW	259				
Qy	253	LRRLRDDDDLETGSGPGGLKGRYFWHVREEFVCEPPLITQHTKLLVLEGOAATLKCKA	312				

Db 260 LRLRLAREDDLEACAPALGGFYFVAVGEEEFVCEPPVTVTHRSPLAVPAGRPAAALRCRA 319  
Qy 313 IGDPSPLIHWAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTIAANAAGEATAMV 372  
Db 320 VGDPEPRVRWVSPQGRLLGNSSRAAFPNGTLELLVTEPDGCGIFTCTIAANAAGEATAAV 379  
Qy 373 EVSI--VOLPHLSNSTRTAPKSRSLSDITGSSKTSRGGGSGGGEPPKSPPERAVLVE 430  
Db 380 ELTVGPPPPQLANSTSCDPPRDGPDALTTPESAASASAKVADTG----PPTDRGVQVTE 435  
Qy 431 VTTTSALVWKSVSAPRVKMYQLQVNCSDDEVLIYRMI PASNKAFVNNLVSGTGYDLC 490  
Db 436 HGATAALVQWPDQRPICIRMYQIYNSSADDILVYRMI PASRSFLTDLASGRTYDLC 495  
Qy 491 VLAMWDDTATLTATNIVGCAOFFTKADYPOCOSMHSOILGCTMILVGGIIVATLLVFI 550  
Db 496 VLAVVGSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGCTMIILGGVIVASVLVFI 555  
Qy 551 VILVRYKVCNHEAP--SKMAAAVSNVYQTNCAQPPPPSSAPAGAPP 596  
Db 556 FVLLMYRYKHGGQPPGKAKIPAPVSSVCSQTNALGPTPTAPPAPPEP 603

RESULT 2  
US-10-245-103-100  
; Sequence 100, Application US/10245103  
; Publication No. US20030068778A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3630R1C112  
; CURRENT APPLICATION NUMBER: US/10/245,103  
; CURRENT FILING DATE: 2002-09-17  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 100  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-245-103-100

Query Match 38.2%; Score 1563.5; DB 9; Length 627;  
Best Local Similarity 53.1%; Pred. No. 1.7e-78;  
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;  
Qy 16 AVVDACPKYCVQNLSIESLGLTCLPCSKGLLPVPPDIDRRTVELRLGNGFIHISQDPANM 75  
Db 23 ATPSPCRRRCRCQTQSPLSVLCFGAGLLFVFPFSLDRRAAELRLADNFIAVSRRRDIAM 82  
Qy 76 TGLVDLTLSNTTSHIQPPFPLDLESRLSHLDNSRLPSLGEDTLRGLVNLQHLIVNNQ 135  
Db 83 TGLHLUSLSNTRIRHVAAGAFADRLRALHLDGNRLTSLGEGQLRGLVNLRLHLSNNQ 142  
Qy 136 LGGIADFAFDFLLTLEDDLSYNNLHCLPWSVRRMVNLHQLSDHNLHLDHIAEGTFAD 195  
Db 143 LAALAGALDDCAETLEDLDSYNNLQPLPWEALGRGNVNTLGLDHNLLASV--PGAFSR 201  
Qy 196 LOKLARLDLTSNRLOKLPDPPIPARSOASALTATPPAPPLS---FSFGNPLHCNCELLW 252  
Db 202 LHKLARLDMTSNRLTTPPPDFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELLW 258  
Qy 253 LRLRLERDDLETGSGPGGLKGRYFHWVREBEFVCEPPLITQHTHKLVLLEGQAATLKCKA 312  
Db 259 LRLRLAREDDLEACASPPALGGRYFWAVGEEFVCEPVPVTHRSPLAVPAGRPAAALRCRA 318  
Qy 313 IGDPSPLIHWAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTIAANAAGEATAMV 372  
Db 319 VGDPEPRVRWVSPQGRLLGNSSRAAFPNGTLELLVTEPDGCGIFTCTIAANAAGEATAAV 378  
Qy 373 EVSI--VOLPHLSNSTRTAPKSRSLSDITGSSKTSRGGGSGGGEPPKSPPERAVLVE 430  
Db 379 ELTVGPPPPQLANSTSCDPPRDGPDALTTPESAASASAKVADTG----PPTDRGVQVTE 434  
Qy 431 VTTTSALVWKSVSAPRVKMYQLQVNCSDDEVLIYRMI PASNKAFVNNLVSGTGYDLC 490  
Db 436 HGATAALVQWPDQRPICIRMYQIYNSSADDILVYRMI PASRSFLTDLASGRTYDLC 494  
Qy 491 VLAMWDDTATLTATNIVGCAOFFTKADYPOCOSMHSOILGCTMILVGGIIVATLLVFI 550  
Db 496 VLAVVGSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGCTMIILGGVIVASVLVFI 554  
Qy 551 VILVRYKVCNHEAP--SKMAAAVSNVYQTNCAQPPPPSSAPAGAPP 596  
Db 556 FVLLMYRYKHGGQPPGKAKIPAPVSSVCSQTNALGPTPTAPPAPPEP 602

RESULT 3  
US-10-245-107-100  
; Sequence 100, Application US/10245107  
; Publication No. US20030068779A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3630R1C71  
; CURRENT APPLICATION NUMBER: US/10/245,107  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046

```

; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-245-107-100

Query Match      38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY 16 AVVDACPKYCVQNLSGLTLCPSKGLFVPPDIDRRTVELRGLGNFIHISRODFANN 75
DB 23 ATPSPCPRRCRCOTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIAVRRDLANN 82
QY 76 TGLVDLTLSRNTISHIQPFSDLESRLSHLDSNRLPSLGEDTTLRGLVNLQHLIVNNQ 135
DB 83 TGLLHLSLSRNTIRHVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSNNQ 142
QY 136 LGGIADAEFEDFLTLEDLDLSYNNLHGLPWDSVRMWNHLSLDHLLDHIAGETPAD 195
DB 143 LAALAGALDDCAETLEDLDLSYNNLEQLPWEALGRGNVNTGLDHNLLASV-PCAFSR 201
QY 196 LOKLARLDLTSNRLOKLPDPDIFARSOASALTATFPAPPLS---FSFGGNPLHNCCELLW 252
DB 202 LHKLARLDWTSNRLTTPDPLFSR---LPLARPRGSPASALVAFGPNPLHNCCELLW 258
QY 253 LRLRERDDDLTCGSPGGKGRYFWHVREBEFVCEPPLITQHTKLLVLEGOAATLKCA 312
DB 259 LRLAREDDLEACASPPALGGRYFWAVGEEFVCEPPVTVTHRSPPPLAVPAGRPAALRCRA 318
QY 313 IGDPSPLTHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
DB 319 VGDPEPRVRWVSPQGRLLGNSSRAFAFNGTLELLVTFEPDGGGIFTCIAANAAGEATAAV 378
QY 373 EVSI--VQLPHLSNSTSRTPPKSRLSDITGSKTSRGGGSGGGGEPKSPPERAVLSE 430
DB 379 ELTVGPPPPQLANSTSCDPPDGDGDDALTPSAASAKVADTG----PPTDRGVQVTE 434
QY 431 VTTTSALVKWSVKSAPRYMYQLQVNCSDDEVLIRYMPASNKAFVNNLVSGTYDLC 490
DB 435 HGATAALVQWPQRPPIGRIMYQIQVNSSADDILVYRMPAESRSFLLTDLASGRTYDLC 494
QY 491 VLAMWDDTATLTATNIVCCAOFFKADYPQCSQHSQILGCTMILVIGGIIVATLLVFI 550
DB 495 VLAVYEDSATGTATRPVGCARFSTEPALRPCAPHAPFLGTMIALGSGVIVASVLYFI 554
QY 551 VILMYRYKVCNHEAP--SKWAAVSNVYQTNQAOPPPSPSSAGAPAGAPP 596
DB 555 FVLLMYRYKHGQPPGKAKIPAPVSVCSQTNALGPTPTAPPAPPEP 602

RESULT 4
US-10-245-143-100
; Sequence 100, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630RIC90
; CURRENT APPLICATION NUMBER: US/10/245,143
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-245-143-100
```

```

Query Match      38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY 16 AVVDACPKYCVQNLSGLTLCPSKGLFVPPDIDRRTVELRGLGNFIHISRODFANN 75
DB 23 ATPSPCPRRCRCOTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIAVRRDLANN 82
QY 76 TGLVDLTLSRNTISHIQPFSDLESRLSHLDSNRLPSLGEDTTLRGLVNLQHLIVNNQ 135
DB 83 TGLLHLSLSRNTIRHVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSNNQ 142
QY 136 LGGIADAEFEDFLTLEDLDLSYNNLHGLPWDSVRMWNHLSLDHLLDHIAGETPAD 195
DB 143 LAALAGALDDCAETLEDLDLSYNNLEQLPWEALGRGNVNTGLDHNLLASV-PCAFSR 201
QY 196 LOKLARLDLTSNRLOKLPDPDIFARSOASALTATFPAPPLS---FSFGGNPLHNCCELLW 252
DB 202 LHKLARLDWTSNRLTTPDPLFSR---LPLARPRGSPASALVAFGPNPLHNCCELLW 258
QY 253 LRLRERDDDLTCGSPGGKGRYFWHVREBEFVCEPPLITQHTKLLVLEGOAATLKCA 312
DB 259 LRLAREDDLEACASPPALGGRYFWAVGEEFVCEPPVTVTHRSPPPLAVPAGRPAALRCRA 318
QY 313 IGDPSPLTHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
DB 319 VGDPEPRVRWVSPQGRLLGNSSRAFAFNGTLELLVTFEPDGGGIFTCIAANAAGEATAAV 378
```



```
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-851-100

Query Match          38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY 16 AVVDACPKYCVCONLSESLGTLCPKSLGLFVPPDIDRRTVELRLGNGFIHISROQDFANM 75
Db 23 ATPSPCRRCRCQOTQSLPLSLVCPGAGLLFVPPSLDRRAELRADNFIAASVRRRLANM 82
QY 76 TGLVDLTLSRNTISHIQPFSLDLNLSLHLDNSRLSLGDETLRGLVNLQHLIVNNQ 135
Db 83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRLHILSNQ 142
QY 136 LGGIADAEAFEDFLTLEDLJLSYNNLHGLPWSVRRMVLNHLQSLDHLNLLDHIAGTFAD 195
Db 143 LAALAGALDDCAETLEDLJLSYNNLEQLPWEALGRNGVNTLGLDHNLLASV-PGAFSR 201
QY 196 LOKLARLDLTSNRLOKLPDPPIFARSQASALTATPAPPPLS---FSFGNPLHNCCELLW 252
Db 202 LHKLARLDMTSNRLTTPDPLFSR---LPLLARPGSPASALVLAFCGNPLHNCCELLW 258
QY 253 LRLERDDDLTFCGSPGGLKGRYFHWVREBEFVCEPPLITQHTHKLVLLEGOAATLKCKA 312
Db 259 LRLAREDDLEACASPPALGGRYFWAVGEEFVCEPPVTVHRSPPLAVPAGRPALRCRA 318
QY 313 IGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
Db 319 VGDPEPRVWVSPQGRLLGNSRRAPFNGTLELLVTEPCDGGIFTFCIAANAAGEATAAV 378
QY 373 EVSI--VOLPHLSNSTSRTPPKSRLSDITGSSKTSRGGGGGGGPPKSPPERAVLVSE 430
Db 379 ELTVGPPPPQLANSTSCDPRDGDPTALTTPSAASAKVADTG---PPTDRGVQVTE 434
QY 431 VTTTSALVKWSVKSAPRVKMYQLOYNCSDDVLIYRMI PASNKA FVNNLVSGTGIDLC 490
Db 435 HGATAALVQWPDORPIPGIRMYQIQYNSSADDILVYRMI PAESRSPLTDLASGRTYDLC 494
QY 491 VLAMWDDTATTLTATNIVCAOFFTKADYPCQOSMHSQILGCTMILVIGIIVATLVFI 550
Db 495 VLAVVEDSATGTATRPVCGARFSTEPALRPCAPHAPFLGGTMIITGALGVIVASVLVFI 554
QY 551 VILMRYKYCNHEAP--SKMAAAVSNVYQTNAGQOPPPSSAPAGAPP 596
Db 555 FVLLMRYKYHGGOPCKAKIPAPVSSVCQTNGALGFTTTPAPPAPPEP 602
```

## RESULT 7

```
US-10-245-883-100
; Sequence 100, Application US/10245883
; Publication No. US20030068783A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
```

```
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanbe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630RIC70
; CURRENT APPLICATION NUMBER: US/10/245,883
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-883-100
```

```
Query Match          38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY 16 AVVDACPKYCVCONLSESLGTLCPKSLGLFVPPDIDRRTVELRLGNGFIHISROQDFANM 75
Db 23 ATPSPCRRCRCQOTQSLPLSLVCPGAGLLFVPPSLDRRAELRADNFIAASVRRRLANM 82
QY 76 TGLVDLTLSRNTISHIQPFSLDLNLSLHLDNSRLSLGDETLRGLVNLQHLIVNNQ 135
Db 83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRLHILSNQ 142
QY 136 LGGIADAEAFEDFLTLEDLJLSYNNLHGLPWSVRRMVLNHLQSLDHLNLLDHIAGTFAD 195
Db 143 LAALAGALDDCAETLEDLJLSYNNLEQLPWEALGRNGVNTLGLDHNLLASV-PGAFSR 201
QY 196 LOKLARLDLTSNRLOKLPDPPIFARSQASALTATPAPPPLS---FSFGNPLHNCCELLW 252
Db 202 LHKLARLDMTSNRLTTPDPLFSR---LPLLARPGSPASALVLAFCGNPLHNCCELLW 258
QY 253 LRLERDDDLTFCGSPGGLKGRYFHWVREBEFVCEPPLITQHTHKLVLLEGOAATLKCKA 312
Db 259 LRLAREDDLEACASPPALGGRYFWAVGEEFVCEPPVTVHRSPPLAVPAGRPALRCRA 318
QY 313 IGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
Db 319 VGDPEPRVWVSPQGRLLGNSRRAPFNGTLELLVTEPCDGGIFTFCIAANAAGEATAAV 378
QY 373 EVSI--VOLPHLSNSTSRTPPKSRLSDITGSSKTSRGGGGGGGPPKSPPERAVLVSE 430
Db 379 ELTVGPPPPQLANSTSCDPRDGDPTALTTPSAASAKVADTG---PPTDRGVQVTE 434
QY 431 VTTTSALVKWSVKSAPRVKMYQLOYNCSDDVLIYRMI PASNKA FVNNLVSGTGIDLC 490
Db 435 HGATAALVQWPDORPIPGIRMYQIQYNSSADDILVYRMI PAESRSPLTDLASGRTYDLC 494
```



Tue Jun 10 08:57:39 2003

us-09-831-846-2.rapb

QY 491 VLAMDDTATTITATNIVCGAOFFTKADYPQCSMHSOILGCTMTLVIGGIIVATLLVFI 550  
 Db 495 VLAVYEDSATGUTATRPVGCARFSTEPALRPGCAPHAPFLGSTMIALGGVIVASVLVFI 554  
 QY 551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNAGQPPPPSSAPAGAPP 596  
 Db 555 FVLLMRYKVHGGOPGPKAKIPAPVSSVCSQTNAGALGPTPTTAPPAPPEP 602

RESULT 8

US-10-237-535-100  
 ; Sequence 100, Application US/10237535  
 ; Publication No. US20030073188A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin  
 ; APPLICANT: Baton, Dan  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stephan, Jean-Philippe  
 ; APPLICANT: Watande, Colin  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; APPLICANT: Fong, Sherman  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P3630R1C3  
 ; CURRENT APPLICATION NUMBER: US/10/237,535  
 ; CURRENT FILING DATE: 2002-09-06  
 ; PRIOR APPLICATION NUMBER: 10/197942  
 ; PRIOR FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/059114  
 ; PRIOR FILING DATE: 1997-09-17  
 ; PRIOR APPLICATION NUMBER: 60/063046  
 ; PRIOR FILING DATE: 1997-10-24  
 ; PRIOR APPLICATION NUMBER: 60/065027  
 ; PRIOR FILING DATE: 1997-11-10  
 ; PRIOR APPLICATION NUMBER: 60/079689  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/086478  
 ; PRIOR FILING DATE: 1998-05-22  
 ; PRIOR APPLICATION NUMBER: 60/087607  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090689  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/091358  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/099803  
 ; PRIOR FILING DATE: 1998-09-10  
 ; PRIOR APPLICATION NUMBER: 60/106932  
 ; PRIOR FILING DATE: 1998-11-03  
 ; PRIOR APPLICATION NUMBER: 60/115554  
 ; PRIOR FILING DATE: 1999-01-12  
 ; PRIOR APPLICATION NUMBER: 60/119342  
 ; PRIOR FILING DATE: 1999-02-09  
 ; PRIOR APPLICATION NUMBER: 60/123957  
 ; PRIOR FILING DATE: 1999-03-12  
 ; PRIOR APPLICATION NUMBER: 60/123972  
 ; PRIOR FILING DATE: 1999-03-11  
 ; PRIOR APPLICATION NUMBER: 60/127372  
 ; PRIOR FILING DATE: 1999-04-01  
 ; PRIOR APPLICATION NUMBER: 60/131271  
 ; PRIOR FILING DATE: 1999-04-27  
 ; PRIOR APPLICATION NUMBER: 60/133459

; PRIOR FILING DATE: 1999-05-11  
 ; PRIOR APPLICATION NUMBER: 60/135725  
 ; PRIOR FILING DATE: 1999-05-25  
 ; PRIOR APPLICATION NUMBER: 60/135729  
 ; PRIOR FILING DATE: 1999-05-25  
 ; PRIOR APPLICATION NUMBER: 60/135750  
 ; PRIOR FILING DATE: 1999-05-25  
 ; PRIOR APPLICATION NUMBER: 60/138385  
 ; PRIOR FILING DATE: 1999-06-09  
 ; PRIOR APPLICATION NUMBER: 60/140653  
 ; PRIOR FILING DATE: 1999-06-22  
 ; PRIOR APPLICATION NUMBER: 60/141037  
 ; PRIOR FILING DATE: 1999-06-23  
 ; PRIOR APPLICATION NUMBER: 60/144732  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/144758  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/144790  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/145228  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/145698  
 ; PRIOR FILING DATE: 1999-07-26  
 ; PRIOR APPLICATION NUMBER: 60/146222  
 ; PRIOR FILING DATE: 1999-07-28  
 ; PRIOR APPLICATION NUMBER: 60/146843  
 ; PRIOR FILING DATE: 1999-08-03  
 ; PRIOR APPLICATION NUMBER: 60/148188  
 ; PRIOR FILING DATE: 1999-08-10  
 ; PRIOR APPLICATION NUMBER: 60/148513  
 ; PRIOR FILING DATE: 1999-08-12  
 ; PRIOR APPLICATION NUMBER: 60/149327  
 ; PRIOR FILING DATE: 1999-08-17  
 ; PRIOR APPLICATION NUMBER: 60/149395  
 ; PRIOR FILING DATE: 1999-08-17  
 ; PRIOR APPLICATION NUMBER: 60/150114  
 ; PRIOR FILING DATE: 1999-08-20  
 ; PRIOR APPLICATION NUMBER: 60/151700  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: 60/151734  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: 60/162506  
 ; PRIOR FILING DATE: 1999-10-29  
 ; PRIOR APPLICATION NUMBER: 60/170262  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: 60/177118  
 ; PRIOR FILING DATE: 2000-01-20  
 ; PRIOR APPLICATION NUMBER: 60/179851  
 ; PRIOR FILING DATE: 2000-02-02  
 ; PRIOR APPLICATION NUMBER: 60/180921  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/187202  
 ; PRIOR FILING DATE: 2000-03-03  
 ; PRIOR APPLICATION NUMBER: 60/198587  
 ; PRIOR FILING DATE: 2000-04-18  
 ; PRIOR APPLICATION NUMBER: 60/199614  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 60/206330  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/206368  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/209832  
 ; PRIOR FILING DATE: 2000-06-05  
 ; PRIOR APPLICATION NUMBER: 60/218371  
 ; PRIOR FILING DATE: 2000-07-13  
 ; PRIOR APPLICATION NUMBER: 60/222695  
 ; PRIOR FILING DATE: 2000-08-02  
 ; PRIOR APPLICATION NUMBER: 60/229896  
 ; PRIOR FILING DATE: 2000-09-01  
 ; PRIOR APPLICATION NUMBER: 60/230621  
 ; PRIOR FILING DATE: 2000-09-05  
 ; PRIOR APPLICATION NUMBER: 60/232887  
 ; PRIOR FILING DATE: 2000-09-15



Tue Jun 10 08:57:39 2003

us-09-831-846-2.rapb

;;  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/091358  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/099803  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/106932  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: 60/115554  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/119342  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: 60/123957  
; PRIOR FILING DATE: 1999-03-12  
; PRIOR APPLICATION NUMBER: 60/123972  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: 60/127372  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/131271  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/133459  
; PRIOR FILING DATE: 1999-05-11  
; PRIOR APPLICATION NUMBER: 60/135725  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 60/135729  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 60/135750  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 60/138385  
; PRIOR FILING DATE: 1999-06-09  
; PRIOR APPLICATION NUMBER: 60/140653  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/141037  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 60/144732  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/144758  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/144790  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145228  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: 60/146222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: 60/146843  
; PRIOR FILING DATE: 1999-08-03  
; PRIOR APPLICATION NUMBER: 60/148188  
; PRIOR FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/148513  
; PRIOR FILING DATE: 1999-08-12  
; PRIOR APPLICATION NUMBER: 60/149327  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/149395  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/150114  
; PRIOR FILING DATE: 1999-08-20  
; PRIOR APPLICATION NUMBER: 60/151700  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/151734  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/162506  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: 60/170262  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 60/177118  
; PRIOR FILING DATE: 2000-01-20  
; PRIOR APPLICATION NUMBER: 60/179851  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 60/180921  
; PRIOR FILING DATE: 2000-02-08  
; PRIOR APPLICATION NUMBER: 60/187202  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: 60/198587  
; PRIOR FILING DATE: 2000-04-18  
; PRIOR APPLICATION NUMBER: 60/199614  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 60/206330  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/206368  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/209832  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: 60/218371  
; PRIOR FILING DATE: 2000-07-13  
; PRIOR APPLICATION NUMBER: 60/222695  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: 60/229896  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/230621  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/232887  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 60/235147  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/261878  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: 60/261910  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/261939  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/262150  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/264395  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/266421  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/267623  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/274399  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 60/280982  
; PRIOR FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: 60/282129  
; PRIOR FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: 60/282199  
; PRIOR FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: 60/290589  
; PRIOR FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: 09/180997  
; PRIOR FILING DATE: 1998-11-19  
; PRIOR APPLICATION NUMBER: 09/267213  
; PRIOR FILING DATE: 1999-03-12  
; PRIOR APPLICATION NUMBER: 09/380137  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380138  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/403297  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 09/423741  
; PRIOR FILING DATE: 1999-11-10  
; PRIOR APPLICATION NUMBER: 09/709238  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 09/802706  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 09/872035  
; PRIOR FILING DATE: 2001-06-01

```
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/924419
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/927796
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/929404
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 09/931836
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/941992
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/946374
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 10/001054
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/081056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 10/119480
; PRIOR FILING DATE: 2002-04-09

Query Match      38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY 16 AVVDACPKYCVQCNLSLGLTCLPSKGLLFVPPDIDRRTVLRLGNGFIHISRODFANM 75
DB 23 ATPSPCPRRCQOTQSLPSVLCPGAGLLFVPPSLDRRAAEURLADNFIAVRRDLANN 82
QY 76 TGLVDLTLSRNTISHIQPFSLDLRLSLHLDNSRLPSLGBDTHLGLVNLQHLIINNQQ 135
DB 83 TGLLHLSLRNTIRHVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSNNQ 142
QY 136 LGGIADAEAFEDLLTLEDLDSYNNLHGLPMDSVRRMVLHQLSDHNLDDHIAEGTPAD 195
DB 143 LAALAGALDDCAETLEDLDSYNNLEQLPWEALGRGNVNTLGLDHNLLASV-PGAFSR 201
QY 196 LOKLARLDLTSNRLQKLPDPFIARSOASALTATPPAPPLS---FSFGNPLHNCCELLW 252
DB 202 LHKLARLDMTSNRLTTPDPPLFSR---LPLARPRGSPASALVLAFCGNPLHNCCELLW 258
QY 253 LRLRLERDDLETGSGPGGLKGRYFWHVRBEEFVCEPPLITQHTHKLVLLEGOAATLKCA 312
DB 259 LRLRLAREDDLEACASPPALGGRYFWAVGEEFVCEPPVTVHRSPPPLAVPAGRPAALRCRA 318
QY 313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGETAMV 372
DB 319 VGDPEPRVRVSPQGRLLGNSSRAAFPNGLTELLVTEFGDGGIFTCIAANAAGETAATV 378
QY 373 EVSI--VQLPHLSNSTSRTPPKSRSLSDITGSKTSRGGGSGGSGGPPKSPPERAVLVEE 430
DB 379 ELTVGPPPPQLANSTSCDPPRDPDALTTPSAASASAKVADTG---PPTDRGVQVTE 434
QY 431 VTTTSALVKWSKSAPRVMYQLOVNCSDDEVLIIYRMI-PAGNKAFFVNNVNLVSGTGIDLC 490
DB 435 HGATAALVQWPQORPIPGIRMYQIQVNSSADDILVYRMI-PAESRSLTDLASGRTYDLC 494
QY 491 VLAMWDDTATTTATNIVCAOFFTKADYPQOSMHSQILGTMILVIGIIVATLLVFI 550
DB 495 VLAVYEDSATGTATRPVGCARFSTEPALRPGCAPHFLGGMFLGGMIALGSGVIVASVLVFI 554
QY 551 VILMYRYKVCNHEAP--SKMAAVSNVYSOTNGAOPPPFPSSPAGAPP 596
DB 555 FVLLMYRYKVGQPGPKAKIPAPVSVCSQTNALGPTTTPAPPAP 602
```

RESULT 10

US-10-238-283-100  
; Sequence 100, Application US/10238283  
; Publication No. US20030073190A1  
; GENERAL INFORMATION:

```
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630RIC15
; CURRENT APPLICATION NUMBER: US/10/238,283
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-238-283-100
```

```
Query Match      38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY 16 AVVDACPKYCVQCNLSLGLTCLPSKGLLFVPPDIDRRTVLRLGNGFIHISRODFANM 75
DB 23 ATPSPCPRRCQOTQSLPSVLCPGAGLLFVPPSLDRRAAEURLADNFIAVRRDLANN 82
QY 76 TGLVDLTLSRNTISHIQPFSLDLRLSLHLDNSRLPSLGBDTHLGLVNLQHLIINNQQ 135
DB 83 TGLLHLSLRNTIRHVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSNNQ 142
QY 136 LGGIADAEAFEDLLTLEDLDSYNNLHGLPMDSVRRMVLHQLSDHNLDDHIAEGTPAD 195
DB 143 LAALAGALDDCAETLEDLDSYNNLEQLPWEALGRGNVNTLGLDHNLLASV-PGAFSR 201
QY 196 LOKLARLDLTSNRLQKLPDPFIARSOASALTATPPAPPLS---FSFGNPLHNCCELLW 252
DB 202 LHKLARLDMTSNRLTTPDPPLFSR---LPLARPRGSPASALVLAFCGNPLHNCCELLW 258
QY 253 LRLRLERDDLETGSGPGGLKGRYFWHVRBEEFVCEPPLITQHTHKLVLLEGOAATLKCA 312
DB 259 LRLRLAREDDLEACASPPALGGRYFWAVGEEFVCEPPVTVHRSPPPLAVPAGRPAALRCRA 318
QY 313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGETAMV 372
DB 319 VGDPEPRVRVSPQGRLLGNSSRAAFPNGLTELLVTEFGDGGIFTCIAANAAGETAATV 378
```





QY 491 VILMWDDTATLTATNIVGCAOFFTKADYPOQOSMHSQILGGTMILVIGGIIVATLLVFI 550  
Db 495 VLVVVEDSAGTATRPVGCARFSTEPALRPGCAPHAPFLGGTMIALGGVIVASVLVFI 554  
QY 551 VILVMRYKVCNHEAP--SKMAAASNVYQTNGAQPPPPSSAPAGAPP 596  
Db 555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCQTNALGPTTTPAPPAP 602

RESULT 14  
US-10-245-730-100  
; Sequence 100, Application US/10245730  
; Publication No. US200300731941  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3630R1C85  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR FILING DATE: 1998-06-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 100  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-245-730-100

Query Match 38.2%; Score 1563.5; DB 9; Length 627;  
Best Local Similarity 53.1%; Pred. No. 1.7e-78;  
Matches 312; Conservative 89; Mismatches 173; Indels 15; Gaps 6;

QY 16 AVVDACPKYCVONLSGSLTCLPSKGLLPVPPDIDRRTVELRGLGNFIHHSRQDFANM 75  
Db 23 ATPSPCPRRCRCQTSQSLPSVLCFGAGLLFPVPSLDRRAELRADNFIAFVRRRLANM 82  
QY 76 TGLVDLTLSNNTTSHIQPFQFSLDESIRSLHDSNRLPSIGEDTTLGLVNLQHLVNNNQ 135  
Db 83 TGLHLSLSNNTTRHVAAGAFADRALRALHLDGNRTSLGEGQLRGLVNLRLHLSNNQ 142  
QY 136 LGGIADAEFDFLLTLEDLDSYNNLHGLPWDSVRRMVLNHLQSLDHNLDHTAEGTFAD 195

Db 143 LAALAAAGALDDCAETLEDLDSYNNLEQLPWEALGRGNVNTLGLDHNLLASV-PGAFSR 201  
QY 196 LOKLARLDITSNRLQKLPPOPIFARSOASALTATPFAPPLS---FSGGNPLHNCCELLW 252  
Db 202 LHKLARLDMTSNRLTTPPPDFSR---LPLARPRGSPASALVLAPOGNPLHNCCELLW 258  
QY 253 LRRLEDDDLTETCGSPGLKGRYFHWVREBEFVCEPPLIQTHTKLLVLEGOAATLKCKA 312  
Db 259 LRLAREDDLEACASPPALGGRYFWAVEEVEFVCEPVPVTHRSPLAVPAGRPAALRCRA 318  
QY 313 IGDPSPLIHWVAPDDRVLVGNSSRTAVYDNGTLDITFTTSQDSGAPTCTIAANAAGATAMV 372  
Db 319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPPGGGIFTCTIAANAAGATAAV 378  
QY 373 EVSI--VOLPHLSNSTRTAPPKSRLSDTICSSKTSRGGGGGGGCEPKSPERAVLYSE 430  
Db 379 ELTVGPPPPQLANSTCDPPRGGDDALTPPSAASAKVADTG---PPTDRGVQVTE 434  
QY 431 VTTTSALVKWSVKSAPRVKMYQLQYNGSDDEVLIYRMI PASNAFVYVNNLVSGTYDLC 490  
Db 435 HGATAALVQWPDQRPICIRMYQIQYNSADDILVYRMI PAESRSFLLTDLASGRTYDLC 494  
QY 491 VILMWDDTATLTATNIVGCAOFFTKADYPOQOSMHSQILGGTMILVIGGIIVATLLVFI 550  
Db 495 VLVVVEDSAGTATRPVGCARFSTEPALRPGCAPHAPFLGGTMIALGGVIVASVLVFI 554  
QY 551 VILVMRYKVCNHEAP--SKMAAASNVYQTNGAQPPPPSSAPAGAPP 596  
Db 555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCQTNALGPTTTPAPPAP 602

RESULT 15  
US-10-245-739-100  
; Sequence 100, Application US/10245739  
; Publication No. US20030073195A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3630R1C86  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR FILING DATE: 1998-06-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 100  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-245-739-100



Search completed: June 8, 2003, 21:38:55  
Job time : 59 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 21:20:53 ; Search time 44 Seconds  
(without alignments)  
1723.865 Million cell updates/sec

Title: US-09-831-846-2  
Perfect score: 4094  
Sequence: 1 METLLGLLAFGMFAVVDA.....DLVGARTGSGSEWWMSTV 789  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	380.5	9.3	707	JC7763	neutonal leucine-r
2	368	9.0	1091	A58532	glial cell membran
3	350	8.5	421	T46266	hypothetical prote
4	286.5	7.0	1531	T42218	slit-1 protein hom
5	263	6.4	1469	B36665	slit protein 2 pre
6	263	6.4	1480	A36665	slit protein 1 pre
7	259	6.3	1523	T13953	MEG5 protein - ra
8	254	6.2	789	T28714	hypothetical prote
9	254	6.2	1355	T28715	hypothetical prote
10	248	6.1	603	JC1282	insulin-like growt
11	241.5	5.9	605	A41915	insulin-like growt
12	237	5.8	1025	T42626	secreted leucine-r
13	236	5.8	361	A53860	chondroaderin pre
14	235.5	5.8	605	JC5239	insulin-like growt
15	232.5	5.7	1535	S46224	peroxidasin - frui
16	231	5.6	603	JC5128	insulin-like growt
17	226.5	5.5	1328	T23007	hypothetical prote
18	221	5.4	907	JG0176	orphan G protein-c
19	208	5.1	907	JG0193	G protein-coupled
20	208	5.1	1896	T08851	Down syndrome cell
21	207.5	5.1	1066	T15864	hypothetical prote
22	203.5	5.0	738	T13938	hypothetical prote
23	203	5.0	821	S06943	brain-derived neur
24	202.5	4.9	1389	T13852	gene wheeler prote
25	202	4.9	333	T34555	hypothetical prote
26	201.5	4.9	420	A53531	oncofetal trophobl
27	201	4.9	626	1 NBUHIA	platelet glycoprot
28	201	4.9	1385	T13887	tir protein - frui
29	198	4.8	536	A34901	lysine carboxypept

30 197 4.8 560 2 A50164 platelet membrane  
31 196 4.8 821 1 A39667 brain-derived neur  
32 195.5 4.8 458 2 T19941 hypothetical prote  
33 195.5 4.8 1344 2 T14316 rig-1 protein - mo  
34 195 4.8 360 2 S06280 decorin precursor  
35 195 4.8 1097 2 A29943 Toll protein precu  
36 193 4.7 562 2 T34319 hypothetical prote  
37 193 4.7 653 2 T25194 hypothetical prote  
38 191 4.7 702 2 T21148 hypothetical prote  
39 190.5 4.7 961 2 T23395 hypothetical prote  
40 189.5 4.6 476 1 A35104 brain-derived neur  
41 188.5 4.6 1257 2 A88536 protein B0523.5 [i  
42 186.5 4.6 375 2 S05390 fibromodulin precu  
43 186.5 4.6 1134 1 A29944 chaptotin precursor  
44 185.5 4.5 354 2 A54544 decorin precursor  
45 185.5 4.5 369 2 S20811 proteoglycan I - m

ALIGNMENTS

RESULT 1

JC7763  
neutonal leucine-rich repeat protein-3 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: JC7763  
R:Fukamachi, K.; Matsuo, Y.; Kitano, C.; Kuchino, Y.; Tsuda, H.  
Biochem. Biophys. Res. Commun. 287, 257-263, 2001  
A:Title: Rat neuronal leucine-rich repeat protein-3: Cloning and regulation of the gene  
A:Reference number: JC7763; PMID:11549284  
A:Contents: Fibrosarcoma cells  
A:Accession: JC7763  
A:Molecule type: mRNA  
A:Residues: 1-707 <FUK>  
A:Cross-references: GB:AF291437  
C:Comment: This protein, a new member of the neuronal leucine-rich repeat protein family  
in protein-protein interaction and functions as a cell adhesion molecule or soluble liga  
C:Genetics:  
A:Gene: nlrr-3  
C:Keywords: cell adhesion

Query Match 9.3%; Score 380.5; DB 2; Length 707;  
Best Local Similarity 22.4%; Pred. No. 3.5e-17;  
Matches 154; Conservative 97; Mismatches 233; Indels 203; Gaps 27;  
QY 12 GMAF-AVVDA-----CPKYCVCO-----NLSESLGTL-CPKGLLVFPDIDRRTV 55  
DB 13 GLAITALVQAGDKKVDPCQCTCEIRPWETPRSYWEASTVDCNDLGLNLFARLPADTQ 72  
QY 56 ELRLGNGFIHISRO-DF-ANMTGLVDLTLSRNTSHIQPFSLDLESLSLHLSNRLP 113  
DB 73 ILLLOTNRIAREHSTDFVNLG---LDLSQNNLSVTNINVQKMSQLSVYLEENKLT 129  
QY 114 SIGETTLRGLVNLQHLIYNNOLGGADEAF----- 144  
DB 130 ELPEKLYGLSLNOELVYNNLLSAPGAFVGLHLLRLHLSNRLQWINSKWEALPN 189  
QY 145 -----EDFLLTLEDL-D-LSY-NNLH 162  
DB 190 LEILMLGNPILRIKMMNFQPLLKURSLVIAGINLTVEPDDALVGLNDSISFYDNRNLN 249  
QY 163 GLPMDSVRRMVMNLHQLSDHLLDHIAGTFA-----DIQ 197  
DB 250 KVPQVALQKAVNLKFLDLNKNPINIRGDSFNMHLKELGINNMPELVSDLSVDNLP 309  
QY 198 KLARLDLTSL-RLQKLPDPDIPARSOASALTATPFA-----PPL-SFSPGGNP 243  
DB 310 DLRKIEATNPRLSYIHENAFRLPKLESMLNSALSALYHGTIESLPNLKEISHSNP 369  
QY 244 LHCNCELLWLR-----RLERDDDLTCTGSPGLKGRYPWHVREEF-----VCPPLIT 292  
DB 370 IRCDVIRWINNKNKTIRFMPEDSL-FCVDPPEFOGQ---NVRQVHFRDMMEICLPLIAP 425

QY 293 QHTKLLVLEGOA-ATLKCAIGDPSLIHWAPD-DRLVGNSSRTA--VYDNGTLDIFI 348  
DB 426 ESFPILOVEADYSVSLHCRATAEQPELYWIITPGSKRLPNTLREKFFVHSEGLDING 485  
QY 349 TTSQDSGAFTCIAANAAGEATAMVEYSIVQLPHLSNSTSRTPPKSRLSDITGSSKTSRG 408  
DB 486 ITPKEGLTYCTIATNLVGADLAKSIMKV----- 513  
QY 409 GGGSGGEBPPKPPERAVLVSEVTTTSALVKSUV-SKAPRVKMYQLQVNCSDDEVLYR 467  
DB 514 ----GGFVPQDNNGSLNIRDIRANSVLVSKANSKILKSSVKWTAFVKTEDSQAQSA 569  
QY 468 MIPASNKAFVNVNLVSGTYDLCVLAMWDDTATLTATNIVCAQOFFTKA---DYPOCOS 524  
DB 570 RIPSQVNVNLTHLKPSTBYKICI-----DIPTIYQSRKOCVNVNTHKSLHDKENK 623  
QY 525 MHSQILGGTMILVIGGI--IVATLWVF 549  
DB 624 SHT-----VFVACVGGLLGIWMLCF 645

RESULT 2  
A58532  
glial cell membrane glycoprotein LIG-1 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Apr-1997 #sequence\_revision 11-Apr-1997 #text\_change 05-Nov-1999  
C:Accession: A58532  
R:Suzuki, Y.; Sato, N.; Tohyama, M.; Wanaka, A.; Takagi, T.  
J. Biol. Chem. 271, 22522-22527, 1996  
A:Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in  
A:Reference number: A58532, MUID:96394313, PMID:8798419  
A:Accession: A58532  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1091 <SUZ>  
A:Cross-references: GB:D78572; NID:g1545806; PIDN:BAAL1416.1; PID:g1545807  
C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan amino-ter  
F:36-61/Domain: leucine-rich alpha-2-glycoprotein repeat homology <PAH>  
F:71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F:262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F:286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F:310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>  
F:334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>  
F:358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>  
F:385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>  
F:409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>  
F:440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>

Query Match 9.0%; Score 368; DB 2; Length 1091;  
Best Local Similarity 21.3%; Pred. No. 4.2e-16;  
Matches 205; Conservative 92; Mismatches 326; Indels 340; Gaps 33;

QY 30 LSESLGTLCPKSGLL-----FVPPDIDRRTVELRLGNFIHISRPQFANMTGLVDIT 82  
DB 188 LRSRLTLRLSKNRITQLPKAFKLP-----RLTQDLNRRNRIEGLTFQGLDLSLEVL 243  
QY 83 LSNRTLSHTQPSFLDLSRLSHLDSNRLPSLCEGTDLGLVNLQHLIVNNOLGGLADE 142  
DB 244 LQRNNTSRITDGAFLGSLKMHVLEHNSLVNNSGLYGLTALHQLHLSNNSISRIQRD 303  
QY 143 APEDFLTLEDLDSLVNNHGLPWDSVRRMVNLHQLSLDHLNLDHIAEGTFADLOKLARL 202  
DB 304 GW-SFOCKLHLLLSFNNTLRDESLAEISSLISIRLSHNAISHIAEGAFKGLKSLRVL 362  
QY 203 DLTSNRLQKLPPDPITFARSOASALT-ATPFAPPLS-----FSFGGNPLH- 245

DB 363 DLDHNEISGITEDTSGAFTGLDNLNLSKLTLCGNKIKSVAKRAFSGLSEHNLNENAIRS 422  
QY 246 -----CNCELLML-----RRLERDDDDLETGCGSPGGLKGRYFW 277  
DB 423 VQDAFAKMKNLKELYISSESFLCDCQLKWLPPMLMGRMLQAFVTATCAHPESLKQSIF 482  
QY 278 HVREBEFVCE-----PPLITO----- 293  
DB 483 SVLPDSFVCDPFPKPOIITQPETTMVAVGKDIRFTCSAASSSSSPMTFAWKDNEVLANA 542  
QY 294 -----HTHK----- 297  
DB 543 DMENFAHVRAQDGEVMEYTTILHLRHVTFGHEGRYQCIITNHFGSTYSHKARLTNVNLP 602  
QY 298 -----LLVLEGOAATLKCAIGDPSLIHW-----VAPDDELVGN 333  
DB 603 FTKIPHDIAIRTTGTARLECAATGHPNPOIAWQKDDGTFPAARERRMHVMPDD----- 656  
QY 334 SRTAVYDNGTLDIFIITTS---QDSGAFTCIAANAAGEATAMVEYSIVQLPHLSNSTSR 390  
DB 657 -----DVFFITDVKIDDMGVYSCTAQSAGSVSANATLTVLETPSL-----A 698  
QY 391 PKSRLSDITGSSKTSRGSGGGGEBPPKPPERAVLVSEVTTTSALVKSUVSKSAPRVK 450  
DB 699 VPLE-----DRVVTVGE---TVAFQCKATGSPTPRIT 727  
QY 451 MYQLQVNCSDDEVLYIRMI PASNKAFVNVNLVSGTYDLCVLAMWDDTAT-TLTATNIVG 509  
DB 728 WLKGRPLSLTE--RHHFTP-GNQLLVQNV-----MDDAGRYTCENSLPLG 772  
QY 510 CAQPFTRKADY---POCQSMHSQILGGTMILVIGGIIVATLWVFIVILMVRYKVCNHEAPS 566  
DB 773 TERAHSQLSILPTPCRK-DGTTVGIFTIIVVCSIVLTSLVWVCIITRKX----- 823  
QY 567 KMAAASVNVYQTNGAQPPPPSSAPAGAPPOGP-----PKVVVRNB----- 607  
DB 824 -----SEEYSVNTDETIVPDPVPSYLSQGLTSDROETVVRTEGGHQANGHIESNGVC 877  
QY 608 LLDFTASTARASDSSSSSLGSGEAGLGRAPWRIPPSAPR-PKPSLRLRMGAF---AS 662  
DB 878 LRD--PSLFPEVDIHSITTCRQPKLCVGYTRFPWKVTEKADRTAAPHHTAHSGSAVCSDCS 935  
QY 663 LDLSQRKEELDSRTPAGAGTSARGHSDREPLLOPPAARASRLPLPLEGKAKRSH 722  
DB 936 TDTAVHPQVPRDSQGP-GTASSQELRQHDREYSP-----HHPYSGTADGSH 981  
QY 723 SFDMGDFAAAAAGGVVPGYGPCKVSNITKRSLSVNGMLLPPEESDLVGARCTFGSSE 782  
DB 982 TL-----SGGSLYPSNHD-----RILPSLKNKAASADGN-GDSS 1014  
QY 783 WVM 785  
DB 1015 WTL 1017

RESULT 3  
T46266  
hypothetical protein DKF2p761A179.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C:Accession: T46266  
R:Blum, H.; Bauersachs, S.; Newes, H.W.; Cassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23034  
A:Accession: T46266  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-421 <AAA>  
A:Cross-references: EMBL:AL137451  
A:Experimental source: adult amygdala; clone DKF2p761A179  
C:Genetics:  
A:Note: DKF2p761A179.1



QY 295 THKLLVLEQQAATLKCKA 312  
Db 285 -----APMECGA 291

RESULT 6

A36665  
slit protein 1 precursor - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C>Date: 30-Apr-1991 #sequence\_revision 30-Apr-1991 #text\_change 02-Aug-2002  
C:Accession: A36665; A31640; S13523  
R:Rothenberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.  
Genes Dev. 4, 2169-2187, 1990  
A:Title: slit: an extracellular protein necessary for development of midline glia and co  
A:Reference number: A36665; MUID:91099665; PMID:2176636  
A:Accession: A36665  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1480 <ROT>  
A:Cross-references: GB:X53959; NID:98614; PIDN:CRA37910.1; PID:98615  
R:Rothenberg, J.M.; Hartley, D.A.; Walther, Z.; Artavanis-Tsakonas, S.  
Cell 55, 1047-1059, 1988  
A:Title: slit: An EGF-homologous locus of *D. melanogaster* involved in the development of  
A:Reference number: A31640; MUID:89077533; PMID:3144436  
A:Accession: A31640  
A:Molecule type: DNA  
A:Residues: 881-1182, 'G', 1185-1404, 'GT', 1463-1464, 'YHA' <RO2>  
A:Cross-references: GB:M23543; NID:G340939; PID:G514357  
C:Genetics:  
A:Gene: FlyBase:slit  
A:Cross-references: FlyBase:FBgn0003425  
A:Introns: 1351/3  
C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein  
C:Keywords: alternative splicing; growth factor  
F:66-91/Domain: proteoglycan amino-terminal homology <PAH1>  
F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>  
F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>  
F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>  
F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>  
F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>  
F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>  
F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>  
F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR19>  
F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR20>  
F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>  
F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>  
F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR21>  
F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR22>  
F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR23>  
F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR24>  
F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>  
F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>  
F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR25>  
F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR26>  
F:791-814/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR27>  
F:815-838/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR28>  
F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>  
F:1028-1061/Domain: EGF homology <EGF>  
F:1068-1099/Domain: EGF homology <EGF2>  
F:1115-1148/Domain: EGF homology <EGF1>

QY 6 GGLAFGMF -----AVVDA-CPKYVCQNLSESLGTLCPKSGLLFPVPPIDRRITVELR 58  
Db 51 GGLGVGHIPOGGVGIVTEARCFVCSCTGLNVD-----CSHRLGTSVPRKISADVERLE 106

QY 59 LGGNFIHISROQFANMTGLVLTLSRNTISHIQPSFLDLSLSLHSDNRLPSLGED 118  
Db 107 LGNNLTIVYETDFQRLTKRLMLQLTNDQIHTIERNFQDLVSLERLDISNNVITVGR 166  
QY 119 TLRLVNLQHLIVNNQLGGIADAEFELLTLEDLDSYNNLHGLPWSVRRMVLHQL 178  
Db 167 VFKAQSLSLQLDNNQITCLDEHAFKG-LVELEILTNNNLTSLF----- 212  
QY 179 SLDHLLDHIAGETPADLQKLARLDTLSNRLQKLPDPFARFSQASALTATPAPPLSFS 238  
Db 213 ---HNI-----FGGLRLRALRLSDN-----PFA----- 233  
QY 239 FGGNPLHNCCELLWRLRLRD-----DDLETQSPGGLKGRYFWHVRREBFVCEPPLITQH 294  
Db 234 -----CDCHLSWLSRFLSATRLAPYTRCQSPQLKQGNVADLHDQEFKCSG--LTEH 284  
QY 295 THKLLVLEQQAATLKCKA 312  
Db 285 -----APMECGA 291

RESULT 7

TI3953  
MEGF5 protein - rat  
N:Alternate names: slit; protein homolog  
C:Species: *Rattus norvegicus* (Norway rat)  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 16-Aug-2002  
C:Accession: TI3953  
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998  
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A:Reference number: Z14126; MUID:98360089; PMID:9693030  
A:Accession: TI3953  
A>Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1523 <NNA>  
A:Cross-references: EMBL:AB011531; NID:G3449291; PIDN:BAA32461.1; PID:G3449292  
C:Genetics:  
C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein  
Query Match 6.3%; Score 259; DB 2; Length 1523;  
Best Local Similarity 27.3%; Pred. No. 9.9e-09;  
Matches 74; Conservative 35; Mismatches 94; Indels 68; Gaps 4;  
QY 20 ACPKYVCQNLSESLGTLCPKSGLLFPVPPIDRRITVELRGGNFIIHSROQDFANMTGLV 79  
Db 279 SCPASACSSNNIVD-----CRGKGLTEIPANLPEGIVEIRLEQNSIKSIPAGAFIOYKKLK 334  
QY 80 DLTLSRNTISHIQPSFLDLSLSLHSDNRLPSLGEDTLRGLVNLQHLIVNNNQLGGI 139  
Db 335 RIDISKQIADIAPDAFQGLKSLTSLVLYGNKITEIPKGLFDGLVSLQLLLNANKINCL 394  
QY 140 ADEAFEDFELLTLEDLDSYNNLHGLPWSVRRMVLHQLSDHLLDHIAGETPADLQKL 199  
Db 395 RVNTQD-----LQNLNLSLYDNKLTQISKGLFAPLQSI 429  
QY 200 ARLDLTLSNRLQKLPDPFARFSQASALTATPAPPLSFSFGGNPLHNCCELLWRLRLRD 259  
Db 430 QTLHLAQ-----NPFVCDCHLKLWADLYLQD 454  
QY 260 DDLET-----CGSPGGLKGRYFWHVRREBFVC 286  
Db 455 NPITSGARCSSPRRLANKRISQIKSKKFC 485

RESULT 8

T28714  
hypothetical protein T21D12.9a - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T28714  
R:Woessner, J.

submitted to the EMBL Data Library, August 1997  
A:Description: The sequence of *C. elegans* cosmid T21D12.  
A:Reference number: Z20514

A:Accession: T28715  
A:Molecule type: DNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-789 <WOE>

A:Cross-references: EMBL:AF016687; PIDN:AAC48096.1; GSPDB:GN00022; CESP:T21D12.9a  
A:Experimental source: strain Bristol N2; clone T21D12

C:Genetics:

A:Gene: CESP:T21D12.9a

A:Map position: 4

A:Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2

Query Match 6.2%; Score 254; DB 2; Length 789;  
Best Local Similarity 21.3%; Pred. No. 8.5e-09;  
Matches 131; Conservative 75; Mismatches 204; Indels 204; Gaps 25;

```
QY 56 ELRLGCGNFIHSRODFANMTGLVDLTLSRNTISHIQPFSLDLESLSHLDS----- 109
DB 83 KLDLASNSITDGTDFSSFNLTVLTKLARNHITTLNQFSRSLKLESLDLTRNMIREV 142
QY 110 -----NRLPSLGEDTLR-----GLVNLQHLIVNNQGLGGADEAFEDFLTL 151
DB 143 RFLAFNQPSLQNVSLARNVDVRLDDGMFYACEGLKHLNLSNRVQAVT-EGWMEGLTSL 201
QY 152 EDLDLSYN-----NLH-----GLPWDSVRRMVNHLQSLDHLNLDH 187
DB 202 EVLDLSYQIOSFHSSWSHTPKLWLSLHNSRIQSLPSGSFVRLQLEELILSANSIDS 261
QY 188 TAEGTFADLQKLARDLTSNRLQ-----KLPPDPPIFAR 220
DB 262 LHKFALVGMSSLHKLDLSSNTLAVCVEDGAVLYNTSMPLSLRFTNNQLRVIPKRAFER 321
QY 221 SQASA---LTATPFA-----PPLSFS---FGGNPLHCNCELLMLR-----RLERDD 261
DB 322 PPALEELDLDTNPIATIHPEAPEPELKLKLVNNSIILCQCISWLASWIYRLKDKSSI 381
QY 262 LETCGSPGLKGRYFWHREEFVC--EPP--LITQHTKLLVLEGOATLKCAIGDPS 317
DB 382 IAKCSYPPPLADLYVVAIDTANLTCHNDSPRAKIVRQPVVESTLIGERKFTCNVYG-AS 440
QY 318 PL-IHW-----VAPDDRVLGNSRTAVYDNGTLD-----IFITTSQDSGA 356
DB 441 PLSIEWRVMEGQPRVLVODSATFLSINRTAVV-NGTDERELAAALLLDNVAMTDNSE 499
QY 357 FTCTAANAAG-EATAMVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGSGG 415
DB 500 YOCVARNRFGSDFSTHVKLQVYQAPKFT----- 527
QY 416 EPPKSPPERAVLVSEVT-----TTSALVKWSVK-SAPRVKMYQLOYNCSDDDEVLIY 466
DB 528 ---YTPEDMPLLVGQTAKFLCAATGTPPEIKWAFEQIPFPAAEARLLVVTDDHI--- 581
QY 467 RMIPASNKAFVNNLVSGTGYDLCLVAMWDDTATLTATNIVGCAQ-----FFTK 516
DB 582 -----YMN-----VTKEDQGYATCATNVAGTQASANLIVFENFF-- 618
QY 517 ADYPOCQSMHSQIL 530
DB 619 -HYSPSPDLSPMLI 631
```

## RESULT 9

T28715

hypothetical protein T21D12.9b - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T28715

R:Woessner, J.

submitted to the EMBL Data Library, August 1997

A:Description: The sequence of *C. elegans* cosmid T21D12.

A:Reference number: Z20514

A:Accession: T28715  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1355 <WOE>

A:Cross-references: EMBL:AF016687; PIDN:AAC48095.1; GSPDB:GN00022; CESP:T21D12.9b

A:Experimental source: strain Bristol N2; clone T21D12

C:Genetics:

A:Gene: CESP:T21D12.9b

A:Map position: 4

A:Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2; 786/3; 84

Query Match 6.2%; Score 254; DB 2; Length 1355;

Best Local Similarity 21.3%; Pred. No. 1.8e-08;

Matches 131; Conservative 75; Mismatches 204; Indels 204; Gaps 25;

```
QY 56 ELRLGCGNFIHSRODFANMTGLVDLTLSRNTISHIQPFSLDLESLSHLDS----- 109
DB 83 KLDLASNSITDGTDFSSFNLTVLTKLARNHITTLNQFSRSLKLESLDLTRNMIREV 142
QY 110 -----NRLPSLGEDTLR-----GLVNLQHLIVNNQGLGGADEAFEDFLTL 151
DB 143 RFLAFNQPSLQNVSLARNVDVRLDDGMFYACEGLKHLNLSNRVQAVT-EGWMEGLTSL 201
QY 152 EDLDLSYN-----NLH-----GLPWDSVRRMVNHLQSLDHLNLDH 187
DB 202 EVLDLSYQIOSFHSSWSHTPKLWLSLHNSRIQSLPSGSFVRLQLEELILSANSIDS 261
QY 188 TAEGTFADLQKLARDLTSNRLQ-----KLPPDPPIFAR 220
DB 262 LHKFALVGMSSLHKLDLSSNTLAVCVEDGAVLYNTSMPLSLRFTNNQLRVIPKRAFER 321
QY 221 SQASA---LTATPFA-----PPLSFS---FGGNPLHCNCELLMLR-----RLERDD 261
DB 322 PPALEELDLDTNPIATIHPEAPEPELKLKLVNNSIILCQCISWLASWIYRLKDKSSI 381
QY 262 LETCGSPGLKGRYFWHREEFVC--EPP--LITQHTKLLVLEGOATLKCAIGDPS 317
DB 382 IAKCSYPPPLADLYVVAIDTANLTCHNDSPRAKIVRQPVVESTLIGERKFTCNVYG-AS 440
QY 318 PL-IHW-----VAPDDRVLGNSRTAVYDNGTLD-----IFITTSQDSGA 356
DB 441 PLSIEWRVMEGQPRVLVODSATFLSINRTAVV-NGTDERELAAALLLDNVAMTDNSE 499
QY 357 FTCTAANAAG-EATAMVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGSGG 415
DB 500 YOCVARNRFGSDFSTHVKLQVYQAPKFT----- 527
QY 416 EPPKSPPERAVLVSEVT-----TTSALVKWSVK-SAPRVKMYQLOYNCSDDDEVLIY 466
DB 528 ---YTPEDMPLLVGQTAKFLCAATGTPPEIKWAFEQIPFPAAEARLLVVTDDHI--- 581
QY 467 RMIPASNKAFVNNLVSGTGYDLCLVAMWDDTATLTATNIVGCAQ-----FFTK 516
DB 582 -----YMN-----VTKEDQGYATCATNVAGTQASANLIVFENFF-- 618
QY 517 ADYPOCQSMHSQIL 530
DB 619 -HYSPSPDLSPMLI 631
```

## RESULT 10

JC1282

insulin-like growth factor-binding protein acid labile chain precursor - rat

C:Species: *Rattus norvegicus* (Norway rat)

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jul-2000

C:Accession: JC1282

R:Dai, J.; Baxter, R.C.

Biochem. Biophys. Res. Commun. 188, 304-309, 1992

A:Title: Molecular cloning of the acid-labile subunit of the rat insulin-like growth fac

A:Reference number: JC1282; MUID:93038676; PMID:1384485

A:Accession: JC1282

A:Molecule type: mRNA

A:Residues: 1-603 <DAI>



A:Cross-references: GB:S46785; NID:g258002; PIDN:AAB23770.2; PID:g5705934  
A:Experimental source: liver  
A>Note: che authors translated the codon AAG for residue 63 as Arg, AAA for residue 205  
C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-603/Product: insulin-like growth factor binding protein, acid labile chain #status  
F:267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 6.1%; Score 248; DB 2; Length 603;  
Best Local Similarity 25.2%; Pred. No. 1.5e-08; Indels 84; Gaps 15;  
Matches 104; Conservative 64; Mismatches 161;

QY 8 LLAFGMAF-----AVVDA-----CPKYCVC--QNLSSEGLTCLPCKGLLFPVDDI 50  
DB 13 LLAFWALGPHLQGTDPGASADAEGPQCPVACTSHDDYTDLSVFCSSKNLTHLPDDI 72  
QY 51 DRRTVELRGGNFIHISQDFANMTGLVDLTLSRNTTSHIOFFSFLDLESRLSHLDN 110  
DB 73 PVSTRALWLDGNNLSIPSAFQNLSSLDFFNLQGSWLSRLEPQALLGLQNLYYLHLERN 132  
QY 111 RLPSIGEDTLRGVNLQHLVNNVNLGGIADAEFDFLLTLEDLDSYNNLHGLPWSVR 170  
DB 133 LRNLAVGLFTTPSLASLSLSLNLGRUEGLFQG-LSHLWDLNLGWSLVVLPDTVFQ 191  
QY 171 RMVNLHOLSLDHNLHDHIAEGTFADLQKLARLDLSNRLOKLPDPPIFARSOASALTATP 230  
DB 192 GLGNLHELVLGNKLYLQALFALFCGLGELRELDLSRNLRSV-----KANVFVHLP 242  
QY 231 FAPPLSFGGNPLHNCNCELLMURLERODDLETGSPGGLKG--RYFW-----HVR----- 280  
DB 243 -----RLQKLYLDRLNLTAVAPGAFGLGMKALRWLDLSHNRVAGL 281  
QY 281 -BEFVCEPRLITQHTHKLVLLEGQAATLKCAKGAIGDPSPLIHWAAPDDRLVGNSSRTAVY 339  
DB 282 MEDTF---PCLGLHV--LRLAHNAIASLRPTFKD-----LHFL--EELQGHNRIRQLG 330  
QY 340 DN-----GLDIF-----ITTSQDSGAFTCIAANAAGATAMVEVSIVQ 378  
DB 331 ERTFEGIGLQLEVLTLNDNQITEVRVGFSGLFNVAVMNLGSLNCLRLSLPERVFQ 383

## RESULT 11

insulin-like growth factor-binding complex acid-labile chain precursor - human  
A:1915  
N:Alternate names: Acid-Labile Subunit (ALS)  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 05-Nov-1999  
C:Accession: A41915  
R:Leong, S.R.; Baxter, R.C.; Camerato, T.; Dai, J.; Wood, W.I.  
Mol. Endocrinol. 6, 870-876, 1992  
A:Title: Structure and functional expression of the acid-labile subunit of the insulin-  
A:Reference number: A41915; MUID:92357025; PMID:1379671  
A:Accession: A41915  
A>Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-605 <LEO>  
A:Cross-references: GB:M86826; NID:g184807; PIDN:AAA36047.1; PID:g184808  
A:Experimental source: liver  
A>Note: sequence extracted from NCBI backbone (NCBIP:110171)

C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology  
F:75-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:99-122/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:123-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:147-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:171-194/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:195-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:243-266/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F:267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F:291-314/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F:315-338/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>  
F:339-362/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>  
F:363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F:387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>  
F:411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>  
F:435-458/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>  
F:459-482/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>  
F:483-506/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>  
F:507-529/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR19>

Query Match 5.9%; Score 241.5; DB 2; Length 605;  
Best Local Similarity 29.7%; Pred. No. 3.9e-08;  
Matches 76; Conservative 41; Mismatches 106; Indels 33; Gaps 4;

QY 20 ACPKYCVC--QNLSSEGLTCLPCKGLLFPVDDIDRRVELRLGNNFIHISRODFANMTG 77  
DB 40 ACPAACVCSDDDADELSVFCSSRNLRTRLPDGVGGTQALWLDGNNLUSSVPPAFAFQLSS 99  
QY 78 LVDLTLSRNTTSHIOFFSFLDLESRLSHLDNRLPSLGEDTLRGLVNLQHLIYNNOLG 137  
DB 100 LGFNLQGGQLSGLEPQALLGLENLCHLERNOLRSALGTFAHTPALASGLSNNRLS 159  
QY 138 GIADAEFDFLLTLEDLDSYNNLHGLPWSVRVMNLHOLSLDHNLHDHIAEGTFADLQ 197  
DB 160 RLEDGLFEG-LGSLWDLNLGWSLAVLPDAAFRGLGSLRELIVLAGNRLAVLQPALFSGLA 218  
QY 198 KLARLDLTSNRLOKLPDPPIFARSOASALTATPAPPPLSFSFGNPLHNCCELLWLRLLE 257  
DB 219 ELRELDLSRNALR-----AIRANVFV-----QLPRLQKLY 248  
QY 258 RDDDLETCTGSPGGLKG 273  
DB 249 LDRNLIAAVAPGAFGL 264

## RESULT 12

T42626  
secreted leucine-rich repeat-containing protein SLIT2 - mouse (fragment)

N:Alternate names: neurogenic extracellular slit protein

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 16-Aug-2002

C:Accession: T42626

R:Holmes, G.P.; Negus, K.; Burridge, L.; Raman, S.; Algar, E.; Yamada, T.; Little, M.H.

Mech. Dev. 79, 57-72, 1998

A:Title: Distinct but overlapping expression patterns of two vertebrate slit homologs in

A:Reference number: 222177; MUID:99279238; PMID:10349621

A:Accession: T42626

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1025 <HOL>

A:Cross-references: EMBL:AF074960; NID:g4151258; PID:g4151259; PIDN:AAD04345.1

C:Genetics:

A:Gene: Slit2

C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein

Query Match 5.8%; Score 237; DB 2; Length 1025;

Best Local Similarity 20.4%; Pred. No. 1.6e-07;

Matches 86; Conservative 60; Mismatches 110; Indels 166; Gaps 10;

QY 20 ACPKYCVCQNLSSEGLTCLPCKGLLFPVDDIDRRVELRL----- 59  
DB 1 ACPEKCRCE-----GTTVDCSNQNLKIPDHIPQYTAEIRLNNEFTVLEATGIFKKLP 54  
QY 60 -----CGNFIHISRODFANMTGLVDLTLSRNTTSHIOFFSFLDLESRLSHLDNRL 112  
DB 55 QLRXINFNNKNTDIEGAFCAGSAGVNEILLTSNLENVQHKMFKGLSKTLMRSNRI 114  
QY 113 PSLGEDTLRGLVNLQHLVNNVNLGGIADAEFDFLLTLEDLDSYN-----NLH----- 162  
DB 115 SCVGNDSFGLGSLVLLSYDNQITTVAPGAF-DXLHSLSTLNLLANPFCNCHLAWLGE 173  
QY 163 ----- 162  
DB 174 WLRRKRIVTGNPRCKQPYFLKEIPIDVAIQDFTCDGNDNDSCLSPSRCPSECTCLDTX 233  
QY 163 -----GLPWSVRVMNLHQ-----LSLDHNLHDHIAEGTF 193

Db 234 VRCSNKGKLVKPGIPKDVTELYDGNQFTLVPKELSNYKHITLTLDLNNRISTLSNQXF 293  
QY 194 ADQKQLARLDLTSNRLOKLPPD-----PIFARSOASALTATPPA 232  
Db 294 SNWTQLLTLILSYNRLRCPPTFDGLSKSLRLSLHNGDISVVPEGAFNDLSALS----- 348  
QY 233 PPLSFSFGNPHUCNCCELLWL-----RRLERDDLETGCGSPGKLGKRGYFWMHREBEFVCEP 288  
Db 349 ---HLAIGANPLYCDNMQWLSDWVKSEYKEPGIARCAGPGMGEMADKLLITTPSKKFTCQG 405  
QY 289 PL 290  
Db 406 PM 407  
RESULT 13  
A53860  
chondroadherin precursor - bovine  
N:Alternate names: 38K leucine-rich protein  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 05-Nov-1999  
C:Accession: A53860  
R:Neame, P.J.; Sommarin, Y.; Boynton, R.E.; Heinegard, D.  
J. Biol. Chem. 269, 21547-21554, 1994  
A:Title: The structure of a 38-kDa leucine-rich protein (chondroadherin) isolated from B  
A:Reference number: A53860; MUID:94342341; PMID:8063792  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-361 <NEA>  
A:Cross-references: GB:U08018; NID:9470671; PIDN:AAA21330.1; PID:9470672  
C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan carboxyl-  
C:Keywords: disulfide bond  
F:300-346/Domain: proteoglycan carboxyl-terminal homology <PCH>  
Query Match 5.8%; Score 236; DB 2; Length 361;  
Best Local Similarity 25.4%; Pred. No. 4.4e-08;  
Matches 90; Conservative 42; Mismatches 122; Indels 100; Gaps 11;  
QY 7 GLLAFGMAFVVDACPKYCVQCNLSGLTCLPSKGLLPVPPDIDRTVELLGGNFIIH 66  
Db 13 GLLA---SLLPALAACPNQCHCH--SDLOHVICDKVGLQKI-PKVSEKTKLLNLQNRNPFV 67  
QY 67 ISRODFANMTGLVD-----LTLSRNTISHIQPSFSLDLES 102  
Db 68 LATNSFRAMPNLVSLHLQHCQIREVAGAFRGKQLIYLYLSHNDIRVLRAQAFDDELTEL 127  
QY 103 RSLHLDNRLPSLGEDTTLRGVLNQLHILVNNQGLGIADAFE----- 145  
Db 128 TVLYLDHNKVTLPRLGSLPLVNLFTLQNNKIRELSGAFQGAQKDLRWLYLSNSLS 187  
QY 146 -----DELLTLEDLDSYNNLHGLPWDSVRMWN- 174  
Db 188 LQFGALDDVENIAKFYLDNQLSSVPSAALSRLRVVEELKSHNPLKSPIDNAFOSFGRY 247  
QY 175 LHLQSLDHLNLDHIAEGTFADQLARLDLTNSNRLOKLPPDPIFARSOASALTATPPAPP 234  
Db 248 LETLWLDNTNLKFSGDGAFGLVTTLKHVILENNRNLHOLPSN--FPFDSLETILT----- 300  
QY 235 LSFSGGNPLHNCCELLMWR-----LERDDLETGCGSPGKLGKRGYFWMHREBE 283  
Db 301 -----NNPWKTCQLGRRLWRLEAKTSRPD--ATCASPAPKFRGQ---HIROTD 343  
RESULT 14  
JC5239  
insulin-like growth factor acid-labile chain - baboon  
C:Species: Papio sp. (baboon)  
C:Date: 17-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-May-1997  
C:Accession: JC5239  
R:Delhanty, P.; Baxter, R.C.  
Biochem. Biophys. Res. Commun. 227, 897-902, 1996

A:Title: The cloning and expression of the baboon acid-labile subunit of the insulin-like  
A:Reference number: JC5239; MUID:97040714; PMID:8886027  
A:Contents: liver  
A:Accession: JC5239  
A:Molecule type: mRNA  
A:Residues: 1-605 <DEL>  
C:Comment: This factor is structurally related to proinsulin and have insuline-like meta  
Query Match 5.8%; Score 235.5; DB 2; Length 605;  
Best Local Similarity 30.5%; Pred. No. 9.8e-08;  
Matches 78; Conservative 34; Mismatches 111; Indels 33; Gaps 4;  
QY 20 ACPKYCVQCNLSE--SLGTLCPSKGLLPVPPDIDRTVELLGGNFIIHISRODFANMTG 77  
Db 40 ACPATCACSVDDEVNELSVFCSSRNLTLPDGPICGTQALMLDSNNLSIPPAAFRNLS 99  
QY 78 LVDLTSLRNTISHIQPSFSLDLESRLHSDNRLPSLGEDTLRGVLNQLHILVNNQLG 137  
Db 100 LAFNLQGGQLGSLEPQALLGLENLCHULHERNQURSLAVGTFTPALALLGLSNNRLS 159  
QY 138 GIADAEFEDFLTLEDLDSYNNLHGLPWDSVRMWNLHQLSLDHLNLDHIAEGTFADIQ 197  
Db 160 RLEDGLFEG-LGNLWDLNGLWNSLAVLPDAAFRGGLRELVLAGNLAYLQPALFSGLA 218  
QY 198 KLARLDLTNSNRLOKLPPDPIFARSOASALTATPPAPPLSFSFGGNPLHNCCELLMWRLE 257  
Db 219 ELRELDLSRNALR-----AIKANVFA-----QLPRLOKLY 248  
QY 258 RDDDLTTCGSPGGLKG 273  
Db 249 LDRNLIAAVAPGAFLG 264  
RESULT 15  
S46224  
peroxidase - fruit fly (Drosophila sp.)  
C:Species: Drosophila sp.  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Oct-2000  
C:Accession: S46224  
R:Nelson, R.E.; Fessler, L.I.; Takagi, Y.; Blumberg, B.; Keene, D.R.; Olson, P.F.; Parke  
EMBO J. 13, 3438-3447, 1994  
A:Title: Peroxidase: a novel enzyme-matrix protein of Drosophila development.  
A:Reference number: S46224; MUID:94341255; PMID:8062820  
A:Accession: S46224  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1535 <NEL>  
A:Cross-references: GB:U11052; NID:9531384; PIDN:AAA61568.1; PID:9531385  
C:Superfamily: peroxidase; myeloperoxidase homology; proteoglycan amino-terminal homolo  
F:19-44/Domain: proteoglycan amino-terminal homology <PAH4>  
F:661-1350/Domain: myeloperoxidase homology <MPX>  
Query Match 5.7%; Score 232.5; DB 2; Length 1535;  
Best Local Similarity 23.5%; Pred. No. 5.5e-07;  
Matches 102; Conservative 56; Mismatches 193; Indels 83; Gaps 13;  
QY 1 METLGGLLAFGMAFVVDACPKYCVQCNLSGLTCLPSKGLLPV--POIDRRTVELR 58  
Db 7.MQLLGLLLLLAGGVQSV-YCPAGCTCLERT-----VRCIRAKLSAVKLPQ-DTQTLDLR 60  
QY 59 LGGNFIIHISRODFANMTGLVDLTLSRNTISHIQPSFSLDLESRLHSDNRLPSLGED 118  
Db 61 F--NHIEELPANAFSGLAQLTTLFLNDNELAYLQDGALNGLTALRFVVLNNNRLSRLPAT 118  
QY 119 TLRGVLNQLHILVNNQLGGIADAEFEDFLTLEDLDSYNNLHGLPWDSVRMWNLHQL 178  
Db 119 IFQMRPLEGIFLENNDIWQLPAGLF-DNLPRNLRLIMYNNKLTQLPVDGFRNLNRLKRL 177  
QY 179 SLDHLNLDHIAEGTFADQLARLDLTNSNRLOKLPPDPIFARSOASALTATPPAPPLSFS 238  
Db 178 RLDGNAID----- 185  
QY 239 FGGNPLHNCNEL--LWLRRLRERDDLE-----TCGSPGGLKGRYFWMHREBEFVCEPPL 290

Db	186	-----IDCNCGVSLW-RRWHLDVQRQLVSVISLTCAAPQMLQNGPSSLGEHHFKCAKPO	239
Qy	291	ITQHTHKLLEGOAATLKCKAIGDPSPLIHWVAPDDRL-VGNSSRTAVYDNGTLDIFIT	349
Db	240	FLVAPQDAQVAAGEQVELSCEVTGLHRPQITWMHNTQELGLEEQTAELPSSGSLHRS	299
Qy	350	TSQDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTRTAPPKSRISDITGSSKT-SRG	408
Db	300	DTSDMGIYQCIARNEMGALRSQP-----VRLVNGGNHPLDSPIDARSNQVWADAGTPTHG	355
Qy	409	GGSGGGEPPKSP	422
Db	356	ATPLPSPSPSHSP	369

Search completed: June 8, 2003, 21:31:21  
Job time : 50 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 19:06:53 ; Search time 36 Seconds  
(without alignments)  
909.023 Million cell updates/sec

Title: US-09-831-846-2  
Perfect score: 4094  
Sequence: 1 METLLGLLAFGMAFAVDDA.....DLVGAGTGGSSSEWNESTV 789

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	388.5	9.5	713	1 GAC1_HUMAN	O75325 homo sapien
2	282.5	6.9	646	1 FLR1_HUMAN	Q9nzul homo sapien
3	263	6.4	1480	1 SLIT2_MOUSE	P24014 drosophila
4	248	6.1	603	1 ALS_RAT	P35859 rattus norv
5	244	6.0	359	1 CHAD_HUMAN	O15335 homo sapien
6	242.5	5.9	966	1 Y918_HUMAN	O94991 homo sapien
7	241.5	5.9	605	1 ALS_HUMAN	P35858 homo sapien
8	240.5	5.9	649	1 FLR3_HUMAN	Q9nzu0 homo sapien
9	240.5	5.9	951	1 LGR4_RAT	Q922h4 rattus norv
10	238	5.8	660	1 FLR2_HUMAN	O43155 homo sapien
11	236	5.8	361	1 CHAD_BOVIN	Q27972 bos taurus
12	236	5.8	977	1 Y848_HUMAN	O94933 homo sapien
13	235.5	5.8	605	1 ALS_PAPHA	O02833 papio hamad
14	234.5	5.7	951	1 LGR4_HUMAN	Q9bxb1 homo sapien
15	232.5	5.7	757	1 LGR7_HUMAN	Q9hbx9 homo sapien
16	231	5.6	603	1 ALS_MOUSE	O55226 mus musculu
17	226	5.5	358	1 CHAD_MOUSE	O70210 rattus norv
18	225	5.5	358	1 CHAD_RAT	Q9nmd4 mus musculu
19	224	5.5	373	1 ASPN_MOUSE	O75473 homo sapien
20	221	5.4	907	1 LGR5_HUMAN	O9bxbn1 homo sapien
21	220	5.4	379	1 ASPN_HUMAN	Q91b75 xenopus lae
22	211	5.2	368	1 PGS1_XENLA	O08770 rattus norv
23	211	5.2	567	1 GPV_RAT	O921p4 mus musculu
24	208	5.1	907	1 LGR5_MOUSE	O60469 homo sapien
25	208	5.1	2012	1 DSCA_HUMAN	P15209 mus musculu
26	203	5.0	821	1 TRKB_MOUSE	P07359 homo sapien
27	201	4.9	626	1 GPBA_HUMAN	Q99mb1 mus musculu
28	201	4.9	905	1 TLR3_MOUSE	O08742 mus musculu
29	200.5	4.9	567	1 GPV_MOUSE	P22792 homo sapien
30	198	4.8	536	1 CBP8_HUMAN	P40197 homo sapien
31	197	4.8	560	1 GPV_HUMAN	Q63604 rattus norv
32	196	4.8	821	1 TRKB_RAT	P21793 bos taurus
33	195	4.8	360	1 PGS2_BOVIN	

RESULT 1

ID	GAC1_HUMAN	STANDARD;	PRT;	713 AA.
AC	O75325;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Glioma amplified on chromosome 1 protein precursor.			
GN	GAC1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Glial tumor;			
RX	MEDLINE=98324709; PubMed=9662332;			
RA	Maifoy B., Almeida A., Zhu X.X., Vogt N., Tyagi R., Muleris M.,			
RA	Dutrillaux A.-M., Dutrillaux B., Ross D., Hanash S.;			
RT	"GAC1, a new member of the leucine-rich repeat superfamily on			
RT	chromosome band 1q32.1, is amplified and overexpressed in malignant			
RT	gliomas."			
RL	Oncogene 16:2997-3002(1998).			
CC	SUBCELLULAR LOCATION: Type I membrane protein (Potential).			
CC	TISSUE SPECIFICITY: OVERAMPLIFIED IN MALIGNANT GLIOMAS.			
CC	SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.			
CC	SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; AF030435; AAC39792.1; -			
CC	MIM; 605492; -			
DR	InterPro; IPR003006; Iq MHC.			
DR	InterPro; IPR003598; Iq_C2.			
DR	InterPro; IPR001611; LRR.			
DR	InterPro; IPR000483; LRR_Cterm.			
DR	InterPro; IPR000372; LRR_Nterm.			
DR	InterPro; IPR003592; LRR_out.			
DR	InterPro; IPR003591; LRR_type.			
DR	Pfam; PF00047; Iq; 1.			
DR	Pfam; PF00560; LRR; 9.			
DR	Pfam; PF01462; LRRNT; 1.			
DR	Pfam; PF01463; LRRCT; 1.			
DR	PRINTS; PR00019; LEURICHRPT.			
DR	SMART; SM00408; IGC2; 1.			
DR	SMART; SM00370; LRR; 6.			
DR	SMART; SM00082; LRRCT; 1.			
DR	SMART; SM00013; LRRNT; 1.			
DR	SMART; SM00369; LRR_TYP; 2.			

P08953 drosophila  
Q9xsd9 sus scrofa  
Q9nr97 homo sapien  
Q9tte2 ovnis aries  
O46542 equus cabal  
O91z25 mus musculu  
P51885 mus musculu  
P13605 bos taurus  
P12024 drosophila  
Q29393 canis famil  
P28654 mus musculu  
P21809 bos taurus

KW	Immunoglobulin domain; Transmembrane; Glycoprotein; Repeat;	
KW	Leucine-rich repeat; Signal.	
FT	SIGNAL	1 18
FT	CHAIN	19 713
FT	DOMAIN	19 713
FT	TRANSMEM	631 651
FT	DOMAIN	652 713
FT	REPEAT	92 115
FT	REPEAT	116 139
FT	REPEAT	140 163
FT	REPEAT	165 187
FT	REPEAT	188 211
FT	REPEAT	213 235
FT	REPEAT	236 259
FT	REPEAT	261 283
FT	REPEAT	309 333
FT	REPEAT	334 357
FT	REPEAT	359 385
FT	DOMAIN	438 504
FT	DISULFID	445 497
FT	CARBOHYD	94 94
FT	CARBOHYD	381 381
FT	CARBOHYD	555 555
FT	CARBOHYD	583 583
SQ	SEQUENCE	713 AA; 78798 MW; EC8BC0FD240C9396 CRC64;
Query Match		
Best Local Similarity 9.5%; Score 388.5; DB 1; Length 713;		
Matches 170; Conservative 91; Mismatches 242; Indels 251; Gaps 27;		
QY	21 CPKVCVCO-----NLSESIGTLCPKSGLLFFVPPDIDRTVELRGGNFIHISQD	71
DB	29 CPQCACQIRPWYTRSSVREATTVCNDLFTAVPPALPACTQTLLQSNISVRDQSE	88
QY	72 FANMTGLVDLTLSRNTISHIQPFSLDESLSLHLSNRLPSLGEDTLRLGLNVOHLIV	131
DB	89 LGLYLANLTEDLSQNSFSRDACDFHALPQLLSLHEENQLTRLEDHSPAGLASIQELY	148
QY	132 NNOLGGIADAF-----	144
DB	149 NHNQLYRIAPRAPSGLSNLLRLHLSNLRALDSRWFMPLNLEILMGGKNVDAILDMN	208
QY	145 -----EDFL-----TLEDLSLYNNLHGLPWSVRRMVNHLQSL	180
DB	209 FRPLANRLSLVLAGMNLREISDYALBGLQSLSPYDQNLARVPRALQVPGKFLDL	268
QY	181 DHNLLDHIABGTF-----DLQKLARLDLTSN-RLQKLPP	214
DB	269 NKNPLQRVGFGDFANMLHLKELGLNMMELVSIDKFAVLNLPDLTKLDITNPRLSFTHP	328
QY	215 DPFIARSOASALTATPFA-----PPL-SFSGGNPLHCNCELLW-----LR	254
DB	329 RAFHLLPQMETLMNNAALSAHQQTVESLNLQEVGLHGNPIRCDCVIRWANATGTRVR	388
QY	255 RLERDDDLTCGSPGKLGKRYFWHVEEF-----VCEPPLITQTH--KLVLVLEQQAAT	307
DB	389 FIETPQSTL--CAEPPDLQR--LPVREVFPREMTDHC-LPLISPRSPFSLQVSGESMV	442
QY	308 LKCKATGDSPLIHWVAPDD-RLVGNSS--RTAVYDNGTLDIFITTSQSGAFTCIAANA	364
DB	443 LHCRALAEPEIYVVTYTAGRLTPAHAGRCRCRVPEGTLELRVTAEEAGLYTCVAQNL	502
QY	365 AGEATAMVEVI-----VOLPHLSN-STSRAPPKSLSDITGSSKT	405
DB	503 VGHADTKTVSVVVGALLQQRDEGQGLELRVQETHPHILLSWVTPNTVSNLTWSSAS	562
QY	406 SRGGGSGGGEPPKPPERAVLVSEVTTTSALYKWSVKSAPRVKMYQLQYNCSDDVLI	465
DB	563 SLRQGA-----TALAR-----	574
QY	466 YRMIPASNAFVNNLVSGTYDLCVLAMWDDTATLTATNIVGCAQFETKADYPQCSM	525
DB	575 ---LPRGTHSYNITRLLOATEYWACLQVAFADAHTQL-----ACVWARTK-----BATSC	621

QY	526 HSQILGGTMLVIGGIIVATLLVFIILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAOPP	585
DB	622 H-RALGDRPGLI-----AILALAVLLUAGLAHLGTGQPRKGV-----GGRRPL	665
QY	586 PPS-----SAPA-----GAP-----PQGPVKVVVRN	606
DB	666 PPAWAFMGWGSAPSVRVVSAPLVLPWNPGRKLPRS	699

RESULT 2

FLR1	HUMAN	STANDARD;	PRT;	646 AA.
ID	FLR1	HUMAN		
AC	Q9NZU1			
DT	15-JUN-2002	(Rel. 41, Created)		
DT	15-JUN-2002	(Rel. 41, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Leucine-rich repeat transmembrane protein FLR1 precursor			
DE	(Fibronectin-like domain-containing leucine-rich transmembrane protein			
GN	FLR1			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.; TISSUE SPECIFICITY, AND GYLCOSYLATION.			
RX	MEDLINE=20112755; PubMed=10644439;			
RA	Lacy S.E., Bonnemant C.G., Buzney E.A., Kunkel L.M.;			
RT	"Identification of FLR1, FLR2, and FLR3: a novel family of			
RT	transmembrane leucine-rich repeat proteins.";			
RL	Genomics 62:417-426(1999).			
CC	-!- FUNCTION: May have a function in cell adhesion and/or receptor			
CC	signaling.			
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).			
CC	-!- TISSUE SPECIFICITY: Expressed in kidney and brain.			
CC	-!- PTM: N-glycosylated.			
CC	-!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.			
CC	-!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; AF169675; AAF28459.1; ALT_INIT.			
DR	Genew; HGNC:3760; FLR1.			
DR	MIM; 604806;			
DR	InterPro; IPR003961; FN III.			
DR	InterPro; IPR001611; LRR			
DR	InterPro; IPR000483; LRR_Cterm.			
DR	InterPro; IPR000372; LRR_Nterm.			
DR	InterPro; IPR003592; LRR_out.			
DR	InterPro; IPR003591; LRR_type.			
DR	Pfam; PF00041; fn3; 1.			
DR	Pfam; PF00560; LRR; 7.			
DR	Pfam; PF01462; LRRNT; 1.			
DR	Pfam; PF01463; LRRCT; 1.			
DR	PRINTS; PR00019; LEURICHRPT.			
DR	SMART; SM00370; LRR; 1.			
DR	SMART; SM00082; LRRCT; 1.			
DR	SMART; SM00013; LRRNT; 1.			
DR	SMART; SM00369; LRR_type; 1.			
KW	Transmembrane; Leucine-rich repeat; Repeat; Glycoprotein; Signal.			
FT	SIGNAL	1 20		
FT	CHAIN	21 646		
FT	DOMAIN	21 524		
FT	TRANSMEM	525 545		
FT	DOMAIN	546 646		

EXTRACELLULAR (POTENTIAL).

LEUCINE-RICH REPEAT TRANSMEMBRANE PROTEIN

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

FT REPEAT 52 77 LRR 1.  
FT REPEAT 78 98 LRR 2.  
FT REPEAT 99 121 LRR 3.  
FT REPEAT 123 147 LRR 4.  
FT REPEAT 148 169 LRR 5.  
FT REPEAT 170 192 LRR 6.  
FT REPEAT 194 218 LRR 7.  
FT REPEAT 219 241 LRR 8.  
FT REPEAT 242 264 LRR 9.  
FT REPEAT 265 288 LRR 10.  
FT DOMAIN 407 485 FIBRONECTIN TYPE-III.  
FT CARBOHYD 221 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 277 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 646 AA; 71359 MW; F2BF5DC3A13C92 CRC64;

Query Match 6.9%; Score 282.5; DB 1; Length 646;  
Best Local Similarity 22.2%; Pred. No. 6.4e-11;  
Matches 145; Conservative 84; Mismatches 218; Indels 205; Gaps 27;

Qy 7 GLLAFGMAFVDA--CPKYCVQNLSGLT-CPKGLLVFPDI--DRRTV----- 55  
Db 12 GLIAP--LTEVIDSTCPVCRDN-----GFIYCNDRGLTSIPADIPDDATTLYLQNNQ 64  
Qy 56 -----ELRGGFIIHISRODFANWTG 77  
Db 65 INNAGIPDOLTKVNVQVIYENDLDEFFINLPRSLRELHLQDNVNTIARDSLRIPL 124  
Qy 78 LVDLTLSRNTIS--HIQFSLDLESRL-----SLHLSNRUPS 114  
Db 125 LEKHLDDNSVTSVIEDADFSKQLKLLFLSRNHLSSIPSGLPHTLEERLDDNRIST 184  
Qy 115 LGEDTLRGLVNLQHLIVNNQLGG--IADEAFEDFLTLEDLDSYNNLHGLPMDSVRM 172  
Db 185 IPLHAFKGLNLSRLVLQDNLNLANQRIADDTFSR--LQNLTELSVRNSLAAPLN--LPS 241  
Qy 173 VNLHQLSLDNLHIAEGTTADLOKLARLDLTSNRLOKLP--DPFARSQASALTAT 229  
Db 242 AHLQKLYLQDNAISHIPYNTLAKRELERLDSNNLTLPRGLFDDDLGNLAQ----- 294  
Qy 230 PFAPPLSFGCNPLHNCCELLWRLRLRRDDLETCSGPGGLKGRYFVHVREEPVCEPP 289  
Db 295 -----LLLRNPFCCNMLWLD-----W----- 314  
Qy 290 LITQHTKLLVLEGOAATLKCAIGDPSPLHVVAPDRLVGNSSRTAVYD--NGTLDIFI 348  
Db 315 -----VKARAAVVNRGLWCQGP-----EKVRG-----MAIKDITSEMDECF 351  
Qy 349 TTSQDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTP-----PKSRLS 397  
Db 352 ETGPQGG-----VANAAAKTTA-----SNHASATTPQGSFLTKAKRPGRLP 394  
Qy 398 DITGSSKTSRGGSGGCGEPPKSPERAVLVSEVTTTSALVKWSVKSAPRVQMYQLOYN 457  
Db 395 DSNIDYPNATGDGAK-----TLAIHVKAUTADSIKITWKATLPASSFRLSWLRLG 444  
Qy 458 CSDSEVLIRMPASNKA--FVYNNLVSGTGYDLCLVAMWDDTATLTATNIVGCAQFFTK 516  
Db 445 HSPAVGSITETLVQGDKTEYLLTALEPKSYIICWVTM--ETSNAVVADETPVCAKAETA 502  
Qy 517 ADYPOCQSMHSQILGGTMI--LVIGGII-----VATLLVFVILMVRVYKVCNHEA 564  
Db 503 DSYGPTTTLNQBQNAAGPMASPLAGIIGGAVALVFLVLGVAICWYV--HQA 552

RESULT 3  
SLIT DROME  
ID SLIT DROME STANDARD; PRT; 1480 AA.  
AC P24014;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Slit protein precursor.  
GN SLI.

OS Drosophila melanogaster. (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91099665; PubMed=2176636;  
RA Rothberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;  
RT "Slit: an extracellular protein necessary for development of midline  
RT glia and commissural axon pathways contains both EGF and LRR  
RT domains.";  
RL Genes Dev. 4:2169-2187(1990).  
CC -!- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND  
CC COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR  
CC MATRIX MOLECULES.  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND  
CC EVENTUALLY DISTRIBUTED ALONG THE AXONS.  
CC -!- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).  
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; X53959; CAA37910.1; -;  
DR PIR; A36665; A36665.  
DR HSSP; PF00740; IEDM.  
DR Flybase; FBgn0003425; sli.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000359; Cys\_knot.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_typ.  
DR InterPro; IPR001791; Laminin\_G.  
DR Pfam; PF000007; Cys\_knot; 1.  
DR Pfam; PF00008; EGF; 7.  
DR Pfam; PF00054; laminin\_G; 1.  
DR Pfam; PF00560; LRR; 16.  
DR Pfam; PF01462; LRRNT; 4.  
DR Pfam; PF01463; LRRCT; 4.  
DR SMART; SM00041; CT; 1.  
DR SMART; SM00179; EGF\_Ca; 2.  
DR SMART; SM00001; EGF\_like; 5.  
DR SMART; SM00370; LRR; 4.  
DR SMART; SM00082; LRRCT; 4.  
DR SMART; SM00013; LRRNT; 4.  
DR SMART; SM00369; LRR\_TYP; 9.  
DR SMART; SM00282; LamG; 1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 3.  
DR PROSITE; PS01185; CTCK\_1; 1.  
DR PROSITE; PS01225; CTCK\_2; 1.  
DR PROSITE; PS00022; EGF\_1; 7.  
DR PROSITE; PS01186; EGF\_2; 5.  
DR PROSITE; PS01187; EGF\_Ca; 2.  
DR PROSITE; PS50025; LAM\_G\_DOMAIN; 1.  
KW Neurogenesis; Glycoprotein; Signal; Alternative splicing;  
KW EGF-like domain; Repeat; Leucine-rich repeat.  
FT SIGNAL 1 36  
FT CHAIN 37 1480 SLIT PROTEIN.

FT	REPEAT	99	122	LRR 1.	SQ	SEQUENCE	1480 AA;	165752 MW;	P9D5925FC170B1C3	CRC64;
FT	REPEAT	123	146	LRR 2.		Query Match	6.4%;	Score 263;	DB 1;	Length 1480;
FT	REPEAT	170	170	LRR 3.		Best Local Similarity	28.0%;	Pred. No. 3.7e-09;		
FT	REPEAT	171	194	LRR 4.		Matches	89;	Conservative	39;	Mismatches 102; Indels 88; Gaps 11;
FT	REPEAT	195	218	LRR 5.						
FT	REPEAT	220	246	LRR 6.						
FT	REPEAT	321	344	LRR 7.						
FT	REPEAT	345	368	LRR 8.	QY	6	GGLAFQMAF----	AVVDA-CPKYCVCONLS	ESLGTLCPSKGLLFVPPDIDRRVTVELR	58
FT	REPEAT	369	392	LRR 9.	Db	51	GGIGSVGIHPGGVGVI	TEARPCVSCITGLNVD----	CSHRLTSVPRKISADVERLE	106
FT	REPEAT	394	416	LRR 10.	QY	59	LGNFIHISRODFANWT	GLVLTLSRNTISHIQP	FLDLSRLSHLDSNLP	SLGED 118
FT	REPEAT	417	440	LRR 11.	QY	107	LQGNLTIVYETDFQRL	TKLRLQLTDNQIHTI	ERNSFQDLVSLERL	DISNNVITVGR 166
FT	REPEAT	522	545	LRR 12.	Db	119	TLRGLVNLQHLIYNN	NOLGGIADEAFDLT	LEDLDSYNNLHGLP	WDSVRRMVNLHQL 178
FT	REPEAT	546	569	LRR 13.	QY	167	VFKGAQSLRSLQDNN	QITCLDEHAFKG-L	VEILEILTNNNLTS	LP----- 212
FT	REPEAT	570	593	LRR 14.	Db	179	SLDHNLLDHAEGT	FADLQKLRLDLTS	NRLOKLPPIFARSO	ASALTATPAPPLSFS 238
FT	REPEAT	618	641	LRR 15.	QY	213	-----HNI-----	FGGLGRRLRLSDN	-----PFA-----	233
FT	REPEAT	643	666	LRR 16.	Db	239	FGNPLHNCCELLWL	RLRLRD-----	DDLETGSPGGLKGR	YFWHVREEEFVCEPLITOH 294
FT	REPEAT	678	701	LRR 17.	QY	234	-----CDCHLSWLS	RFLRSATRLAPYTR	CQSPSQLKQNVADL	HDQEFKCSG--LIEH 284
FT	REPEAT	720	743	LRR 18.	QY	295	THKLLVLEGOAATL	KCKA 312		
FT	REPEAT	745	764	LRR 19.	Db	285	-----APMECGA	291		
FT	REPEAT	765	788	LRR 20.						
FT	REPEAT	790	812	LRR 21.						
FT	REPEAT	813	836	LRR 22.						
FT	REPEAT	838	861	LRR 23.						
FT	REPEAT	907	944	LRR 24.						
FT	DOMAIN	946	983	EGF-LIKE 1.						
FT	DOMAIN	985	1022	EGF-LIKE 2.						
FT	DOMAIN	1024	1062	EGF-LIKE 3.						
FT	DOMAIN	1064	1100	EGF-LIKE 4.						
FT	DOMAIN	1111	1149	EGF-LIKE 5.						
FT	DOMAIN	1152	1325	EGF-LIKE 6.						
FT	DOMAIN	1333	1392	EGF-LIKE 7.						
FT	DOMAIN	1409	1480	CRCK.						
FT	CARBOHYD	111	111	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	207	207	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	357	357	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	435	435	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	783	783	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	788	788	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	788	788	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	958	958	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	998	998	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	1060	1060	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	1159	1159	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	1175	1175	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	1243	1243	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	CARBOHYD	1292	1292	N-LINKED (GLCNAC. . .)						(POTENTIAL).
FT	DISULFID	911	922	BY SIMILARITY.						
FT	DISULFID	916	932	BY SIMILARITY.						
FT	DISULFID	934	943	BY SIMILARITY.						
FT	DISULFID	950	961	BY SIMILARITY.						
FT	DISULFID	955	971	BY SIMILARITY.						
FT	DISULFID	973	982	BY SIMILARITY.						
FT	DISULFID	989	1001	BY SIMILARITY.						
FT	DISULFID	995	1010	BY SIMILARITY.						
FT	DISULFID	1012	1021	BY SIMILARITY.						
FT	DISULFID	1028	1041	BY SIMILARITY.						
FT	DISULFID	1035	1050	BY SIMILARITY.						
FT	DISULFID	1052	1061	BY SIMILARITY.						
FT	DISULFID	1068	1079	BY SIMILARITY.						
FT	DISULFID	1073	1088	BY SIMILARITY.						
FT	DISULFID	1090	1099	BY SIMILARITY.						
FT	DISULFID	1115	1125	BY SIMILARITY.						
FT	DISULFID	1120	1137	BY SIMILARITY.						
FT	DISULFID	1139	1148	BY SIMILARITY.						
FT	DISULFID	1357	1368	BY SIMILARITY.						
FT	DISULFID	1362	1380	BY SIMILARITY.						
FT	DISULFID	1382	1391	BY SIMILARITY.						
FT	DISULFID	1409	1443	BY SIMILARITY.						
FT	DISULFID	1423	1457	BY SIMILARITY.						
FT	DISULFID	1434	1473	BY SIMILARITY.						
FT	DISULFID	1438	1475	BY SIMILARITY.						
FT	DISULFID	1442	1479	BY SIMILARITY.						
FT	VARSPIC	1394	1404	MISSING (IN SHORT ISOFORM).						



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC EMBL; S46785; AAB23770.2; .
CC PIR; JCI282; JCI282.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000483; LRR Cterm.
CC InterPro; IPR00372; LRR Nterm.
CC InterPro; IPR003592; LRR out.
CC InterPro; IPR003591; LRR typ.
CC Pfam; PF00560; LRR; 18.
CC Pfam; PF01462; LRRNT; 1.
CC Pfam; PF01463; LRRCT; 1.
CC PRINTS; PRO0019; LEURICHRPT.
CC SMART; SM00370; LRR; 5.
CC SMART; SM00082; LRRCT; 1.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00369; LRR typ; 9.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 603
FT REPEAT 52 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 146 168
FT REPEAT 169 192
FT REPEAT 194 216
FT REPEAT 217 240
FT REPEAT 242 264
FT REPEAT 266 288
FT REPEAT 289 312
FT REPEAT 313 336
FT REPEAT 337 360
FT REPEAT 361 384
FT REPEAT 386 408
FT REPEAT 409 432
FT REPEAT 433 456
FT REPEAT 458 478
FT REPEAT 479 504
FT REPEAT 506 529
FT REPEAT 543 566
FT CARBOHYD 64 64
FT CARBOHYD 85 85
FT CARBOHYD 96 96
FT CARBOHYD 368 368
FT CARBOHYD 515 515
FT CARBOHYD 578 578
FT CARBOHYD 586 586
SQ SEQUENCE 603 AA; 66811 MW; DCD7637D94A5037C CRC64;

Query Match 6.1%; Score 248; DB 1; Length 603;
Best Local Similarity 25.2%; Pred. No. 1e-08;
Matches 104; Conservative 64; Mismatches 161; Indels 84; Gaps 15;

QY 8 LLAFCMAF-----AVVDA---CPKYVCV--QNLSESLTCLPSKGLFPVDPDI 50
DB 13 LLAFVWALGPCHQGTDPGASADAEQPCVACTCSHDDYTDLSVFCSSKNLTHLPDDI 72
QY 51 DRRTVELRGGNFIHISQDFANMTGLVDLTLSRTTISHQIPFSFLDLESLSRLSDSN 110
DB 73 PVSTRALWLDGNLSSISPAFQNLSSLDLFLNQLQSWLSLEPQALLGLQNLVYLHLEN 132
QY 111 RLPSLGEEDTLRLGNVLQHLVNNNOLGGIADAFEDFLTLLTBDLDSYNNLHGLPWSVR 170
DB 133 RLRLNLAUGLFTHTPSLASLSLSSNLLGRLEGLFQG-LSHLWDLNLGWSLVLPDVTVFQ 191
QY 171 RMVNLHQLSDHNLDDHDAEGTFADLOKLARDLTSNLOKLPPDPPIFARQASALTATP 230
DB 192 GLGNLHELVLGNKLTLYLPALFCGLBELRELDLSRNALRSV-----KANVFVHLP 242
```

## RESULT 5

CHAD\_HUMAN

ID CHAD\_HUMAN STANDARD; PRT; 359 AA.

AC O15335; Q96RJ5;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Chondroadherin precursor (Cartilage leucine-rich protein).

GN CHAD.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98008928; PubMed=9344663;

RA Grover J., Chen X.-N., Korenberg J.R., Roughley P.J.;

RT "The structure and chromosome location of the human chondroadherin

RL gene (CHAD).";

RL Genomics 45:379-385(1997).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21413956; PubMed=11445564;

RA Maansson B., Wenglen C., Moergelin M., Saxne T., Heinegaard D.;

RT "Association of chondroadherin with collagen type II.";

RL J. Biol. Chem. 276:32883-32888(2001).

CC !- FUNCTION: Promotes attachment of chondrocytes, fibroblasts, and

CC osteoblasts. This binding is mediated (at least for chondrocytes

CC and fibroblasts) by the integrin alpha(2)beta(1). May play an

CC important role in the regulation of chondrocyte growth and

CC proliferation (By similarity).

CC !- SUBUNIT: Mostly monomeric (By similarity). Interacts with collagen

CC type II.

CC !- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).

CC !- TISSUE SPECIFICITY: Present in chondrocytes at all ages.

CC !- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCAN

CC (SLRPS) FAMILY. CLASS IV SUBFAMILY.

CC !- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; U96769; AAC13410.1;

DR EMBL; U96767; AAC13410.1; JOINED.

DR EMBL; U96768; AAC13410.1; JOINED.

DR EMBL; AF371328; AAK51556.1; .

DR Genew; HGNC:1909; CHAD.

DR MIM; 602178; .

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR Cterm.

DR InterPro; IPR000172; LRR Nterm.

DR InterPro; IPR003592; LRR out.

DR InterPro; IPR003591; LRR typ.

DR Pfam; PF00560; LRR; 10.



```

322 322 WVA PDDRLVGNSSRTAYVDN-GTLDIFITTS-----QDSGATFC----- 359
349 349 RVRP TSR--QPSKDLGYSNTGPSIAYTKGPVPLECTACSCNLOI SDLGLNVNCQERK 405
360 360 -----IAAN--AAGEATAMVEVSI VQLPHLSNSTSRTAPPKSR-LSD 398
406 406 IESIAELQPKPNPKMYLTENYIAVVRRTDFLEATGLDLHLHGN--NRISMIQDRAFCD 463
399 399 ITGSSKTSRGGGGGGGEPKSPPERAVLVSEVITTSALVKWSVSKSAPR-----VKM 451
464 464 LTNRLRLYLNGN-----RIBELSPELPYGLQSLOQY 493
452 452 YOLOYN-CSDDEVILYIRMI PASNKAFVNNLV---SG--TG YDLCVLAMMDDTATTLTA 504
494 494 LFLQYNIIREIQSGTFDPVPVNLQLLFLLNNLQAMPSPGVFSGTLRLRLNLSNHTSTLPV 553
505 505 TNIVGCAQFFTKADYPQCOSMHSQI LGGTMTILVIGGIIVATLLVFIVILMYRYKVCNHEA 564
554 554 SGVLDQLKSLI QID-----LHDNPWDC TCDIVGNKLWVEQLKGVULDEV---IC--KA 602
565 565 PSKMAAAVSNVYSGTNGAQQPPPPSAPAGAPQGPQPKVVVNRNELLDFTASLARASDSSSS 624
603 603 PKKFAE--TDMRSIKSELLCEPDYSDVVVSTPTPSSIQVPARTSAV---TPAVRLNSTGAP 657
625 625 SSLGSGEAAGLGRAPRIPPSAPRKPSLDLRLMGAFAS-----LDLKSORKEELDSRT- 678
658 658 ASLGAGGGAS-----SVPLSVLILSLLLVFMVSFVAAGLFVLVWKERKKQNSDHTSTN 711
679 679 -----PAGRGAGTSARGHHSDREPLL---GPPAARARSLLPL 714
712 712 NSDVSSFNMQSVYGGGGGTGCHPHAHVHRGPALPKVTKTAPGHVYVEIHPFL 764

```

## RESULT 7

ALS_HUMAN	STANDARD;	PRT;	605 AA.
ID	ALS_HUMAN		
AC	P35858;		
DT	01-JUN-1994 (Rel. 29, Created)		
DT	01-JUN-1994 (Rel. 29, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Insulin-like growth factor binding protein complex acid labile chain precursor [ALS].		
DE	IGFALS OR ALS.		
GN	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
OX	[1]_SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RP	TISSUE=Liver;		
RC	MEDLINE=92357025; PubMed=1379671;		
RC	Leong S.R., Baxter R.C., Camerato T., Dai J., Wood W.I.;		
RT	"Structure and functional expression of the acid-labile subunit of		
RT	the insulin-like growth factor-binding protein complex.";		
RT	Mol. Endocrinol. 6:870-876(1992).		
RL	[2]		
RL	SEQUENCE FROM N.A.		
RP	Frankland J.;		
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.		
RL	[3]		
RP	SEQUENCE OF 28-35.		
RP	MEDLINE=89308584; PubMed=2473065;		
RA	Baxter R.C., Martin J.B., Beniac V.A.;		
RT	"High molecular weight insulin-like growth factor binding protein		
RT	complex. Purification and properties of the acid-labile subunit from		
RT	human serum.";		
RL	J. Biol. Chem. 264:11843-11848(1989).		
CC	-!- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT		
CC	IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.		
CC	-!- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH		
CC	IGF-I OR IGF-II AND IGFBP-3.		
CC	-!- SUBCELLULAR LOCATION: Extracellular.		

QY 138 GIADAFEDFLTLTDLSDSYNNHGLPQSDVRRMVNHLQSLDHLNLDHIAEGTFADLQ 197  
 DB 160 RIEDGLFEG-IGSLWDNLGNLWNSLAVLPDPAFRGIGSLRELVLGNRLAYLQPALFSGLA 218  
 QY 198 KIARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLISFSGGNPLHCNCELLWLRLLE 257  
 DB 219 ELRELDLSRNALR-----AIKANVFV-----QLPRLOKLY 248  
 QY 258 RDDDLTCTGSPGLK 273  
 DB 249 LDRNLIAAVAPGAFILG 264

## RESULT 8

FLR3\_HUMAN  
 ID FLR3\_HUMAN STANDARD; PRT: 649 AA.  
 AC Q9NZ00; Q96KB1; Q9P259; Q96K42; Q96K39;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Leucine-rich repeat transmembrane protein FLRT3 precursor  
 DE (Fibronectin-like domain-containing leucine-rich transmembrane protein  
 3).  
 DE FLRT3 OR KIAA1469.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RX MEDLINE=20112755; PubMed=10644439;  
 RA Lacy S.E., Bonnemant C.G., Buzney E.A., Kunkel L.M.;  
 RT "Identification of FLRT1, FLRT2, and FLRT3: a novel family of  
 RT transmembrane leucine-rich repeat proteins";  
 RL Genomics 62:417-426(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20277482; PubMed=10819331;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XVII.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro";  
 RL DNA Res. 7:143-150(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo, and Teratocarcinoma;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehtvasilaino M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

Marsh V.L., Martin S.L., McConnachie L.J., Mclay K., McMurray A.A.,  
 Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 Palmer K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showken R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Symons A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,  
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Bentley D.R., Beck S.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20";  
 RL Nature 414:865-871(2001).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: May have a function in cell adhesion and/or receptor  
 CC signaling.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).  
 CC -!- TISSUE SPECIFICITY: Expressed in kidney, brain, pancreas, skeletal  
 CC muscle, lung, liver, placenta, and heart.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF169677; AAF28461.1; -  
 DR EMBL; AB040902; BAA95993.1; ALT\_INIT.  
 DR EMBL; AK027297; BAB55023.1; -  
 DR EMBL; AK027670; BAB55282.1; ALT\_INIT.  
 DR EMBL; AK027694; BAB55303.1; ALT\_INIT.  
 DR EMBL; AL1132826; CAB86687.1; -  
 DR EMBL; BC020870; AAB20870.1; -  
 DR Genew; HGNC:3762; FLRT3.  
 DR MIM; 604808; -  
 DR InterPro; IPR003961; FN III.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000483; LRR Cterm.  
 DR InterPro; IPR000372; LRR Nterm.  
 DR InterPro; IPR003592; LRR out.  
 DR InterPro; IPR003591; LRR typ.  
 DR Pfam; PF00041; fn3; 1.  
 DR Pfam; PF00560; LRR; 10.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR Pfam; PF01463; LRRCT; 1.  
 DR SMART; SM00060; FN3; 1.  
 DR SMART; SM00370; LRR; 1.  
 DR SMART; SM00082; LRRCT; 1.  
 DR SMART; SM00013; LRRNT; 1.  
 DR SMART; SM00369; LRR\_TYP; 1.  
 DR Transmembrane; Leucine-rich  
 DR SIGNAL 1 28  
 DR CHAIN 29 649  
 FT repeat; Repeat; Glycoprotein; Signal.  
 FT POTENTIAL.  
 FT LEUCINE-RICH REPEAT TRANSMEMBRANE PROTEIN  
 FT FLRT3  
 FT EXTRACELLULAR (POTENTIAL).  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT LRR 1.  
 FT LRR 2.  
 FT LRR 3.  
 FT LRR 4.  
 FT LRR 5.  
 FT LRR 6.  
 FT LRR 7.  
 FT LRR 8.  
 FT 224

```

FT REPEAT 247 269 LRR 9.
FT REPEAT 270 293 LRR 10.
FT DOMAIN 405 485 FIBRONECTIN TYPE-III.
FT CARBOHYD 226 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 198 198 L -> P (IN REF. 3; BAB55282).
FT CONFLICT 307 307 W -> R (IN REF. 3; BAB55282).
FT CONFLICT 400 400 H -> Q (IN REF. 3; BAB55023).
FT CONFLICT 519 519 K -> Q (IN REF. 3; BAB55023).
FT CONFLICT 638 638 D -> G (IN REF. 3; BAB55282).
SQ SEQUENCE 649 AA; 73003 MW; 9EF666C46181F08 CRC64;

Query Match 5.9%; Score 240.5; DB 1; Length 649;
Best Local Similarity 21.8%; Pred. No. 3.4e-08;
Matches 141; Conservative 90; Mismatches 220; Indels 195; Gaps 28;

QY 3 TLLGGLAFGMAFVADCPKVCVQ-----NLSESLGTL----- 37
Db 13 TKIGLFQVAPUSVMAKSCPSVCRDAGIYCNDRFLTSIPTGIPEDATTLYLQNNQINN 72
QY 38 --CPS--KGLFV-----PPDIDRTVELRLGGNFIHISQDQFANMTGLVD 80
Db 73 AGIPSDLNKLKVERIYLHNSLDEFTNLPKYVKELHLOENNIRITIVDSLSKIPYLEE 132
QY 81 LVLRSNTIS--HIQFSPFLDLSLR-----SLHSDNRLPSLGE 117
Db 133 LHLDDNSVSAVIEGAFRDSNYLRLFLSRNHLSTIPWGLPRTTEELRLDDNRISTISS 192
QY 118 DTLRGVNLHILVNNOLG--GIADAEFEDLLTLEDLDSYNNLHGLPWDSVRMVNL 175
Db 193 PSLOGLTSLKRLVNLGNNHGLGDKVFN--LVNLTELSLVRSNLTAAPVNLPG--TNL 249
QY 176 HOLSDHNLHDIAGTADLOKLARLDLTSNRLKLPDPDIFARSAQASALTATPAPPL 235
Db 250 RKLYLQDNHINRVPNFAFSYLRQLYELDSNNLSNL--PQIF--DDLDNIT----- 298
QY 236 SFSFGNPLHNCCELLWRLRLERDDLET-----CGSPGLKGRYFWHVRBEF 284
Db 299 QLILRNPNWYCGCKMKWR-----DWLQSLPVKVNVRGLMCAPEKVRGMAIKDLNAELF 353
QY 285 VCEP-----PLITQHTHLVLVSGQ--AATLKCAKADPSPLHWPADPDLVGNSSR 335
Db 354 CDKDSGVISTIGITTAIPNTVYPAQGVMPVTPKQPDIKNPK-----LTKDHQHTGSPSR 408
QY 336 TAVYDNGTLDITITSQ-----DSGAFTCIAAN----- 363
Db 409 KTI-----TITVKSVTSDTHISWKLALPMTALRLSWLKLGHSPAGSITETITVGERSEY 464
QY 364 --AAGEATAMVEVSIVQLP-----HLNSNT-----SRTAPPKSRSLSDITGSSKTSRGGGS 412
Db 465 LVTALEPDSPYKVCMPVPMETSNLYLFDPTPVCIETETAPLRMYNPTTTLNREQEK----- 519
QY 413 GGEPEPKSP--PERAVLSEVTTTS-----ALYKWSV-----SKAPRVKMYQ 453
Db 520 ----EPYKNPLPAAIIGGAVLVALVTIALLCVWYVHRNGSLFSRNCAYSKGRRRRDDYA 576
QY 454 LOYNGSDDEVL-----IYRMIPASN-----KAFVNNLVSGTGYDL 489
Db 577 EAGTKKDSNLSIRETSFQMLPISNEPISKEEFVHTIPPNGMNL 622

RESULT 9
LGR4_RAT STANDARD; PRT; 951 AA.
AC Q922H4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
GN GPR48 OR LGR4
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.
RX TISSUE=Ovary;
RX MEDLINE=99065210; PubMed=9849958;
RA Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
RT "Characterization of two LGR genes homologous to gonadotropin and
RT thyrotropin receptors with extracellular leucine-rich repeats and a G
RT protein-coupled, seven-transmembrane region.";
RL Mol. Endocrinol. 12:1830-1845(1998).
CC !- FUNCTION: Orphan receptor.
CC !- SUBCELLULAR LOCATION: Integral membrane protein.
CC !- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC !- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF061443; AAC77910.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000372; LRR_Nterm.
CC InterPro; IPR003592; LRR_Out.
CC InterPro; IPR003591; LRR_Typ.
CC Pfam; PF00001; 7tm_1; 1.
CC Pfam; PF00560; LRR; 15.
CC Pfam; PF01462; LRRNT; 1.
CC PRINTS; PR00237; GPCRRHODOPSIN.
CC PRINTS; PR00019; LEURICHRPT.
CC SMART; SM00370; LRR; 4.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00369; LRR_Typ; 5.
CC PROSITE; PS00237; G-PROTEIN RECP_F1_1; FALSE_NEG.
CC PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Signal; Glycoprotein;
KW Repeat; Leucine-rich repeat.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 951 LEUCINE-RICH REPEAT-CONTAINING G PROTEIN-
FT COUPLED RECEPTOR 4.
FT DOMAIN 25 544 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 545 565 1 (POTENTIAL).
FT DOMAIN 566 575 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 576 596 2 (POTENTIAL).
FT DOMAIN 597 619 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 620 640 3 (POTENTIAL).
FT DOMAIN 641 661 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 662 682 4 (POTENTIAL).
FT DOMAIN 683 703 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 704 724 5 (POTENTIAL).
FT DOMAIN 725 756 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 757 777 6 (POTENTIAL).
FT DOMAIN 778 783 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 784 804 7 (POTENTIAL).
FT DOMAIN 805 951 CYTOPLASMIC (POTENTIAL).
FT REPEAT 55 79 LRR 1.
FT REPEAT 81 103 LRR 2.
FT REPEAT 104 127 LRR 3.
FT REPEAT 128 151 LRR 4.
FT REPEAT 153 175 LRR 5.
FT REPEAT 176 199 LRR 6.
FT REPEAT 200 223 LRR 7.
FT REPEAT 225 247 LRR 8.
FT REPEAT 248 270 LRR 9.
FT REPEAT 272 294 LRR 10.
FT REPEAT 318 341 LRR 11.
FT REPEAT 342 366 LRR 12.
FT REPEAT 367 387 LRR 13.

```

```

FT REPEAT 388 411 LRR 14.
FT REPEAT 413 435 LRR 15.
FT DISULFID 618 693 BY SIMILARITY.
FT CARBOHYD 68 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 951 AA; 104138 MW; EDD56AC072123461 CRC64;

Query Match 5.9%; Score 240.5; DB 1; Length 951;
Best Local Similarity 22.3%; Pred. No. 5.7e-08;
Matches 152; Conservative 84; Mismatches 230; Indels 217; Gaps 29;

QY 7 GLLAF-----GMAFVAVDACPKYVCQNLSESLGTLCPKSGLLFVPP----- 48
DB 6 GLLCFALGLLGSAGPSGAAPLCAAPCSDG---DRRVDCSGKGLTAVPEGLSAFTQAL 62
QY 49 DIDRTV-----ELRGGNFIIHSRQDFANWTGLVDLTLSNTISHIQP 93
DB 63 DISMNNITQLPEDAFKSPFLEELQAGNDLSLIHPKALSGELKVLTLQNNQRTVPS 122
QY 94 FSPDLDESRLHDSNRLPSLGEDTLRLGLVNLQHLIVNNNOLOGGIADAEFDLTLLED 153
DB 123 EAIHGLSALQSRLDANHITSVPEDSFGVLQRLHMLDNDNSITEVPRPLSN-LPTLQA 181
QY 154 LOLSNNHGLPWSVRRMVNLHQLSDHNLHDHTAEGTFADLOKLARDLTSNLOKLP 213
DB 182 LTLALNMISSIPDAFTNLSLVHLHNNKIKLSQHCDFGLDNLDTLDLNNYLDPEP 241
QY 214 PDIFARSOASALTATFAPPLSF-----SFGGNPLHCNCELLWLARLERDDDL 262
DB 242 -----QATKALPSLKELGFHNSISVIPDGAFGNPL-----LRTIHLVD-- 281
QY 263 ETCGSPGLKGRYFMH-----VREBEFVCEPLITQHTH-KLLVLEGOAATLKCA 312
DB 282 ---NPLSFVGNFAFNLSDHLCLVIRGASLVQWFNFNLGTGTHLESILTLTG---TKISS 333
QY 313 IGDPSPLIHWAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANRAGEATMV 372
DB 334 I-----PDD-LCQNKMLR-----TLDSYNNIRDLPSF-----NGCRAL 368
QY 373 EVSIV--QLPHLSNSTSRAPPKSLRSLDITGSSKTSRGGGGGGGPPKSPPERAVLVE 430
DB 369 EISLQRNQISLIKENTFO-GLTSRLDLRSN-----LIRE 403
QY 431 V-----TTTSALYKWSVKSAPR-----VMYQLQVNCSDDEVLIYRMIPASNKAF 476
DB 404 IHSGAFAPAKGTIINLDVSFNELTSFPTEGLNGLNQLKLVGNFKLKDAL-----AARDFA 457
QY 477 VVNNLVSGTYDLCVLAMWDDTATLTATNIVGC---AQFPTKADYPOCOSM----- 525
DB 458 NLRSLSPVYAYQCC--APW-----GCDSYANLNTEDNSPQESHVTKKGAATD 502
QY 526 -----HSQI-----LGGTMI-LVIGGIITVATLLVFIILMV 555
DB 503 AANVTSTAENEHSQIIHCTPSGAFKPCPEYLLGSWMIRLTVMFIFVALLFNLLVILT 562
QY 556 RYKVCNHEAPSKM---AAAVSNV 575
DB 563 VFASCSLSPASKLFIGLISVSNL 585
```

```

RESULT 10
FLR2 HUMAN STANDARD; PRT; 660 AA.
AC O43155;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich repeat transmembrane protein FLRT2 precursor
DE (Fibronectin-like domain-containing leucine-rich transmembrane protein
```

```

DE FLRT2 OR KIAA0405.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
(1)
RP SEQUENCE FROM N.A.; TISSUE SPECIFICITY, AND GLYCOSYLATION.
RX MEDLINE=20112755; PubMed=10644439;
RA Lacy S.E., Bonnemant C.G., Buzney E.A., Kunkel L.M.;
RT "Identification of FLRT1, FLRT2, and FLRT3: a novel family of
RT transmembrane leucine-rich repeat proteins.";
RL Genomics 62:417-426(1999).
(2)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98116655; PubMed=9455477;
RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VIII.
RT 78 new cDNA clones from brain which code for large proteins in
RT vitro.";
RL DNA Res. 4:307-313(1997).
CC -!- FUNCTION: May have a function in cell adhesion and/or receptor
CC signaling.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Expressed in pancreas, skeletal muscle, brain,
CC and heart.
CC -!- PTM: N-glycosylated.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF169676; AAF28460.1; -
CC EMBL; AB007865; BAA23701.1; -
CC Genew; HGNC:3761; FLRT2.
CC MIM; 604807; -
CC InterPro; IPR003961; FN III.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000483; LRR Cterm.
CC InterPro; IPR000372; LRR Nterm.
CC InterPro; IPR003592; LRR Out.
CC InterPro; IPR003591; LRR typ.
CC Pfam; PF00041; fn3; 1.
CC Pfam; PF00560; LRR; 8.
CC Pfam; PF01462; LRRNT; 1.
CC Pfam; PF01463; LRRCT; 1.
CC SMART; SM00370; LRR; 3.
CC SMART; SM00082; LRRCT; 1.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00369; LRR typ; 1.
KW Transmembrane; Leucine-rich repeat; Repeat; Glycoprotein; Signal.
FT SIGNAL 1 35
FT CHAIN 36 660
FT DOMAIN 36 541
FT TRANSMEM 542 562
FT DOMAIN 563 660
FT REPEAT 62 87
FT REPEAT 88 108
FT REPEAT 109 131
FT REPEAT 132 157
FT REPEAT 159 181
FT REPEAT 183 202
FT REPEAT 203 228
FT REPEAT 229 251
```

```

RESULT 11
CHAD_BOVIN
ID _CHAD_BOVIN STANDARD; PRT; 361 AA.
AC Q27972;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chondroadherin precursor (Cartilage leucine-rich protein) (38 kDa bone
DE protein).
DN CHAD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
SEQUENCE FROM N.A.A., AND PARTIAL SEQUENCE.
RP
RP TISSUE=Cartilage;
RC
RX MEDLINE=94342341; PubMed=8063792;
RT Neame P.J., Sommarin Y., Boynton R.E., Heinegaard D.;
RA "The structure of a 38-kDa leucine-rich protein (chondroadherin)
RT isolated from bovine cartilage."
RL J. Biol. Chem. 269:21547-21554(1994).

```

[2]  
SEQUENCE OF 25-55 AND 77-97.  
RP TISSUE=Bone;  
RC MEDLINE=95113864; PubMed=7814406;  
RA Hu B., Coulson L., Moyer B., Price P.A.;  
RX "Isolation and molecular cloning of a novel bone phosphoprotein  
RT related in sequence to the cystatin family of thiol protease  
RT inhibitors.";  
RL J. Biol. Chem. 270:431-436(1995).  
CC -!- FUNCTION: Promotes attachment of chondrocytes, fibroblasts, and  
CC osteoblasts. This binding is mediated (at least for chondrocytes  
CC and fibroblasts) by the integrin alpha(2)beta(1). May play an  
CC important role in the regulation of chondrocyte growth and  
CC proliferation.  
CC -!- SUBUNIT: Mostly monomeric. Interacts with collagen type II (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Extracellular matrix.  
CC -!- TISSUE SPECIFICITY: Cartilage.  
CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCAN  
CC (SLRPS) FAMILY. CLASS IV SUBFAMILY.  
CC -!- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
-----  
DR EMBL; U08018; AAA21330.1; -;  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR\_Typ.  
DR Pfam; PF00560; LRR; 10.  
DR Pfam; PF01462; LRRT; 1.  
DR Pfam; PF01463; LRRT; 1.  
DR SMART; SM00082; LRRT; 1.  
DR SMART; SM00013; LRRT; 1.  
DR SMART; SM00369; LRR\_Typ; 5.  
KW Signal; leucine-rich repeat; Repeat; Glycoprotein;  
KW Extracellular matrix.  
FT SIGNAL 1 24 OR 23 (IN SOME ISOFORM(S)).  
FT CHAIN 25 352 CHONDROADHERIN.  
FT CHAIN 25 361 CHONDROADHERIN, MINOR FORM.  
FT REPEAT 52 75 LRR 1.  
FT REPEAT 76 99 LRR 2.  
FT REPEAT 100 123 LRR 3.  
FT REPEAT 124 147 LRR 4.  
FT REPEAT 149 171 LRR 5.  
FT REPEAT 173 195 LRR 6.  
FT REPEAT 196 219 LRR 7.  
FT REPEAT 220 243 LRR 8.  
FT REPEAT 245 268 LRR 9.  
FT REPEAT 269 292 LRR 10.  
FT REPEAT 294 319 LRR 11.  
FT BY SIMILARITY.  
FT DISULFID 25 40  
FT DISULFID 306 348  
FT DISULFID 308 328  
FT CARBOHYD 146 146 O-LINKED (POTENTIAL).  
FT CONFLICT 25 25 C -> Y (IN REF. 2).  
FT CONFLICT 29 29 C -> W (IN REF. 2).  
FT CONFLICT 31 31 C -> H (IN REF. 2).  
FT CONFLICT 40 40 C -> L (IN REF. 2).  
FT CONFLICT 52 52 S -> R (IN REF. 2).  
SQ SEQUENCE 361 AA; 40884 MW; DA79DC98AD3DD1F8 CRC64;  
  
Query Match 5.8%; Score 236; DB 1; Length 361;  
Best Local Similarity 25.4%; Pred.No. 3e-08;  
Matches 90; Conservative 42; Mismatches 122; Indels 100; Gaps 11;  
  
OV 7 GLLAFGMAFAVWDACPKYCVQNLSSLSGLTCLSPKGLLFFVPPIDIRRTVELRGLGNFIH 66



```
Db 13 GLLA--SLLPALAACPNCHCH--SDLOHVICDKVGLQKI-PKVSEKTKLLNLRNPPV 67
Qy 67 ISRODFANMTGLVD-----LTLRSNTISHIQPFSELDLES 102
Db 68 LATNSFRAMPNVLVSLHLOHCQIREVAAGAFGLKQLIYLYSHNDIRVLRAGAFDDUTEL 127
Qy 103 RSLHLSNRLPSLGBDTLRLGLVNLQHLVNNQGLGGIADAEF----- 145
Db 128 TYLYLDHNKVTLPGLLSPLVNLPLQILNNKIRLSRGAFOGAKDLRWLYLSNSLSS 187
Qy 146 -----DFLLTLEDLSYNNHLGLPWSVRRMVN- 174
Db 188 LQPGALDDVENIAKYLDORNQLSSVPSAALSRLRVEELKLSHPNPKIPDNAPQSGRY 247
Qy 175 LHLQSLDHLNLDHIAEGTFADLQKLRDLTSNRLOKLPDPPIFARSOASALTATPPAPP 234
Db 248 LETLWLDNTNLEKFSGDAFLGVTTLKHVHLENNRLHQLPSN--FPFDSLETTLT----- 300
Qy 235 LSFSGGNPLHNCCELLMRR-----LERDDDLTCTGSPGGLKGRYFWHVRREE 283
Db 301 -----NNPWKCTQLRLRLWLEAKTRPD--ATCASPAPFRGQ---HIRDTD 343

RESULT 12
Y848 HUMAN
ID Y848 HUMAN STANDARD; PRT; 977 AA.
AC O94933;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein KIAA0848.
GN KIAA0848.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RS SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.;"
RL DNA Res. 5:355-364 (1998).
CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AB020655; BAA74871.1;
CC InterPro; IPR001611; LRR.
CC DR InterPro; IPR000483; LRR_Cterm.
CC DR InterPro; IPR000372; LRR_Nterm.
CC DR InterPro; IPR003592; LRR_out.
CC DR InterPro; IPR003591; LRR_typ.
CC Pfam; PF00560; LRR; 9.
CC Pfam; PF01463; LRRCT; 2.
CC SMART; SM00370; LRR; 4.
CC SMART; SM00082; LRRCT; 2.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00369; LRR_TYP; 3.
CC Hypothetical protein; Repeat; Leucine-rich repeat.
KW REPEAT 76
FT REPEAT 99 LRR 1.
FT REPEAT 100 123 LRR 2.
FT REPEAT 125 147 LRR 3.
```

```
FT REPEAT 148 171 LRR 4.
FT REPEAT 172 195 LRR 5.
FT REPEAT 197 219 LRR 6.
FT REPEAT 407 430 LRR 7.
FT REPEAT 431 454 LRR 8.
FT REPEAT 455 478 LRR 9.
FT REPEAT 480 502 LRR 10.
FT REPEAT 503 526 LRR 11.
FT REPEAT 528 550 LRR 12.
FT DOMAIN 723 POLY-GLY.
SQ SEQUENCE 977 AA; 109005 MW; 3C936B7E0003DF54 CRC64;

Query Match 5.8%; Score 236; DB 1; Length 977;
Best Local Similarity 24.8%; Pred. No. 1.2e-07;
Matches 111; Conservative 42; Mismatches 178; Indels 116; Gaps 15;

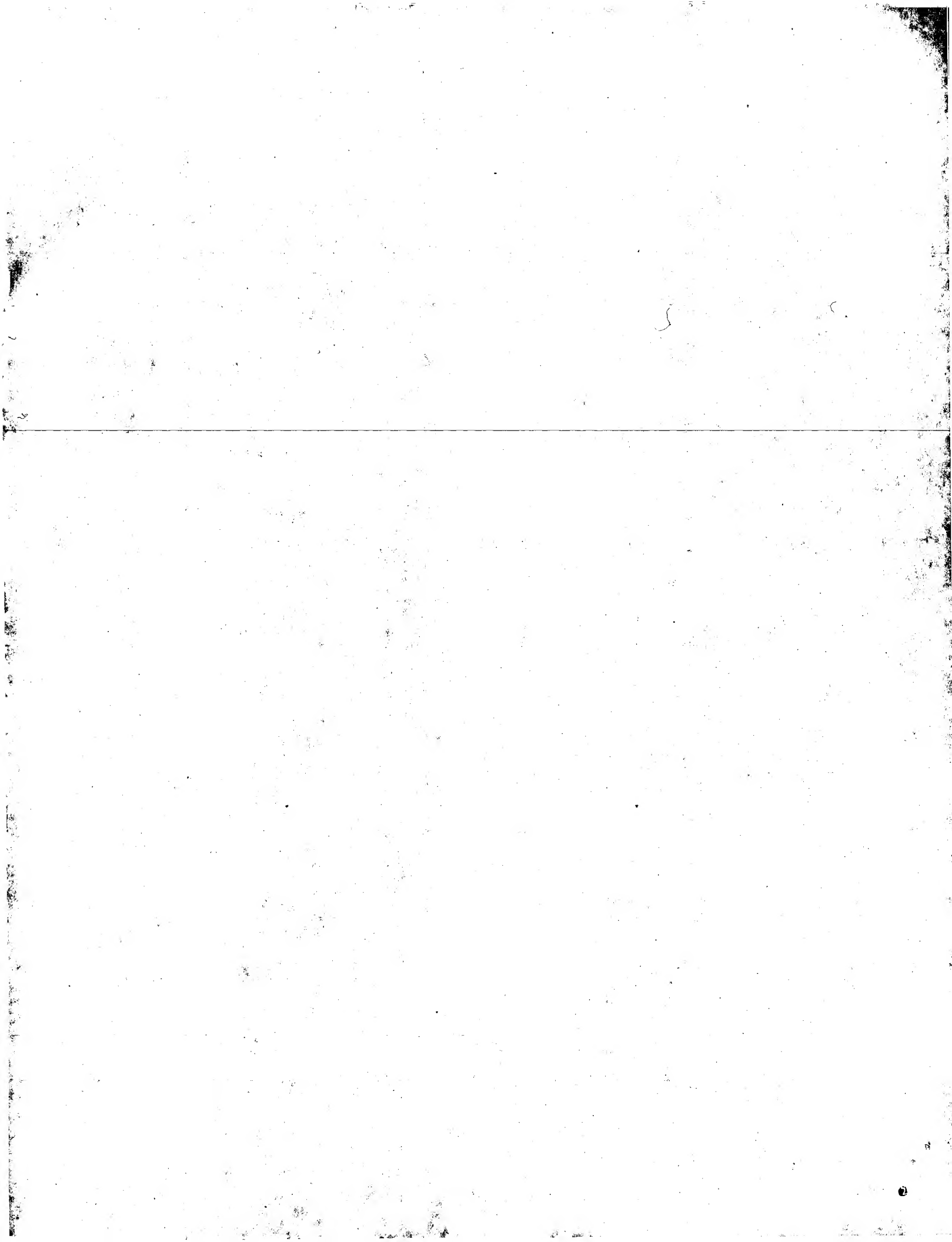
Qy 21 CPKYCVCONLSLGLT--CPSKGLFVPPDIDR--RTVELRLGGNFITHISRODFANMT 76
Db 373 CPTGCTCNLHNDGLTVNCKERGNNISELLRPPLNAKLYLSSNLIQKIYRSDFWNFS 432
Qy 77 GLVDLTLSRNTISHIQPFSELDLESRLSLHLSNRLPSLGEDTLRLGLVNLQHLVNNQGL 136
Db 433 SLDLLHLGNRRISYVQDGAFLNPLNLSFLNGNDIEKLTGPMFRGLQSLHYLYFEFNV 492
Qy 137 GGIADAEFEDLTLTLEDLDSYNNHLGLPWSVRRMVNHLQSLDHLNLDHIAEGTFADL 196
Db 493 REIQPAAFS-----LMPNLKLLFLNNLRLTPTDAFAG- 526
Qy 197 QKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLSFGGNPLHNCCELL-----W 252
Db 527 TSLARLNLKRNLYFLYLPVAGVLEHLNAI-----VOIDNENWDCTCDLVPPKQW 576
Qy 253 LRLRLR---DDDLTCTGSPGGLKGRYFWHVR--EEFVCEPPLITQTHKLLVLEGOAA 306
Db 577 IETISSVVVGDVL--CESPENLTHR---DVRTIEVLVC----- 611
Qy 307 TLKCKAIGDPSPLIHWAP-----DRLVGNSSRTAVYD---NGTLDIFITTSQD 353
Db 612 -----PEMLHVAPAGESPAQFGDHLIGAPTSASPYEFPSPGPPVLSVLISL 660
Qy 354 SGAFTCAANAAGATAMVEYSIVOLPHLNSNSTRTAPPKSRSLSDITGSSKTSR----- 407
Db 661 LVLFSSAVFVAAGLFAYVLRRLRRKKLPFRSK-----RQEGVDLTIGIQCHRLPEDG 712
Qy 408 ---GGGGGGGEGPPKPPPERAVLVSEV 431
Db 713 GGGGGGGGGGRPTLSSPEKAPPVGHV 739

RESULT 13
ALS PAPA
ID ALS PAPA STANDARD; PRT; 605 AA.
AC O02833;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
DE precursor (ALS).
GN IGFALS OR ALS.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9557;
RN 1;
RS SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97040714; PubMed=8986027;
RA Delhanty P., Baxter R.C.;
RT "The cloning and expression of the baboon acid-labile subunit of the
RT insulin-like growth factor binding protein complex.;"
RL Biochem. Biophys. Res. Commun. 227:897-902 (1996).
CC -!- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
```









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 20:50:18 ; Search time 62 Seconds  
(without alignments)  
2622.116 Million cell updates/sec

Title: US-09-831-846-2

Perfect score: 4094

Sequence: 1 METLLGGLLAFMAFVDDA.....DLVGARTGSGSEWVMESTV 789

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteria.\*

17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4094	100.0	832	4 Q9ULH4	Q9ulh4 homo sapien
2	4075	99.5	789	6 Q9BE71	Q9be71 macaca fasc
3	3885.5	94.9	788	11 Q9CYK3	Q9cyk3 mus musculus
4	1821	44.5	719	4 Q9GN16	Q9gn16 homo sapien
5	1573	38.4	628	4 Q9BTD0	Q9btd0 homo sapien
6	1571.5	38.4	700	4 Q9BT24	Q9bt24 homo sapien
7	1320.5	32.3	492	11 Q9BKT6	Q9bkt6 mus musculus
8	1063.5	26.0	542	4 Q9BES9	Q9bes9 homo sapien
9	775	18.9	450	4 Q9EC50	Q9ec50 homo sapien
10	653	16.0	324	4 Q9BWCJ	Q9bcwj homo sapien
11	435.5	10.6	640	4 Q9HCJ2	Q9hcj2 homo sapien
12	418.5	10.2	606	4 Q9BZ20	Q9bz20 homo sapien
13	385	9.4	649	4 Q9A85	Q9a85 homo sapien
14	383.5	9.4	648	11 Q9V135	Q9v135 mus musculus
15	383	9.4	653	4 Q9HBW1	Q9hbwl homo sapien
16	380.5	9.3	707	11 Q9ESY6	Q9esy6 rattus norv

17	380	9.3	614	6 Q9N008	Q9n008 macaca fasc
18	379	9.3	614	4 Q96FES	Q96fes homo sapien
19	374	9.1	614	11 Q9DIT0	Q9difo mus musculus
20	373	9.1	707	11 Q97860	Q97860 mus musculus
21	368	9.0	708	4 Q9H3W5	Q9h3w5 homo sapien
22	368	9.0	1091	11 P70193	P70193 mus musculus
23	365	8.9	1496	4 Q92626	Q92626 homo sapien
24	362.5	8.9	1093	4 Q96JAI	Q96jal homo sapien
25	361.5	8.8	1094	4 Q9BYB8	Q9byb8 homo sapien
26	354.5	8.7	718	13 Q73675	Q73675 xenopus lae
27	350.5	8.6	422	4 Q9NT99	Q9nt99 homo sapien
28	341.5	8.3	892	5 Q9VM16	Q9vm16 drosophila
29	334.5	8.2	730	4 Q9P231	Q9p231 homo sapien
30	331.5	8.1	716	11 Q61809	Q61809 mus musculus
31	322.5	7.9	1065	4 Q94898	Q94898 homo sapien
32	312	7.6	1021	5 Q9V430	Q9v430 drosophila
33	304	7.4	473	6 Q9N0E3	Q9n0e3 macaca fasc
34	303	7.4	473	4 Q9BZ86	Q9bz86 homo sapien
35	296	7.2	441	11 Q99PH1	Q99phi mus musculus
36	295	7.2	473	11 Q99PI8	Q99pi8 mus musculus
37	295	7.2	1316	5 Q9VQ25	Q9vg25 drosophila
38	294	7.2	836	5 Q9V9V6	Q9v9v6 drosophila
39	292.5	7.1	892	5 P91644	P91644 drosophila
40	290.5	7.1	894	5 Q9VKG1	Q9vkg1 drosophila
41	290.5	7.1	1530	11 Q9WUG5	Q9wugs rattus norv
42	286.5	7.0	1531	11 Q88279	Q88279 rattus norv
43	284.5	6.9	1173	5 Q9V7J8	Q9v7j8 drosophila
44	284.5	6.9	1525	4 Q9Y507	Q9y507 homo sapien
45	283.5	6.9	798	4 Q8WWZ2	Q8wwz2 homo sapien

## ALIGNMENTS

RESULT 1

Q9ULH4 PRELIMINARY; PRT; 832 AA.  
ID Q9ULH4;  
DC Q9ULH4;  
AC 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE KIAA1246 protein (Fragment).  
GN KIAA1246.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=20039619; PubMed=10574462;  
RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XV.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins (in vitro).  
RL DNA Res. 6:337-345 (1999).  
DR EMBL; AB033072; BA86560.1;  
DR HSSP; P56276; 1TLK  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR003598; Ig\_C2.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR Pfam; PF00041; fn3; 1.  
DR Pfam; PF00560; LRR; 7.  
DR Pfam; PF01463; LRRCT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00060; FN3; 1.  
DR SMART; SM00408; IGC2; 1.

DR SMART; SM00370; LRR; 1.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 3.  
KW Immunoglobulin domain.  
FT NON\_TER 1  
SQ SEQUENCE 832 AA; 88978 MW; 9B0E80F7BFF974A8 CRC64;  
Query Match 100.0%; Score 4094; DB 4; Length 832;  
Best Local Similarity 100.0%; Pred. No. 5, 4e-310;  
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 METLLGGLLAFGMAFAVVDACPKYCVQNLSGLTLCPSKGLLFVPPDIDRRTVELRLG 60  
Db 44 METLLGGLLAFGMAFAVVDACPKYCVQNLSGLTLCPSKGLLFVPPDIDRRTVELRLG 103  
QY 61 GNFIHISRODFANMTGLVDLTLSRNTISHIQPFSLDLESRLSHLDSNRLPSLGEDTL 120  
Db 104 GNFIHISRODFANMTGLVDLTLSRNTISHIQPFSLDLESRLSHLDSNRLPSLGEDTL 163  
QY 121 RGLVNLQHLI VNNNOLGGIADAEFDLTLEDLDSYNNLHGLPWSVRRMVLHQLSL 180  
Db 164 RGLVNLQHLI VNNNOLGGIADAEFDLTLEDLDSYNNLHGLPWSVRRMVLHQLSL 223  
QY 181 DHNLDDHIAEGTFADLQKLARLDLTLSNRLOKLPDPPIFARSOASALTATPPAPPLSFSG 240  
Db 224 DHNLDDHIAEGTFADLQKLARLDLTLSNRLOKLPDPPIFARSOASALTATPPAPPLSFSG 283  
QY 241 GNPLHNCCELLWLRLERDDDLTCTGSPGGLKGRYFHWVREBEFVCEPPLIQTHTKLLV 300  
Db 284 GNPLHNCCELLWLRLERDDDLTCTGSPGGLKGRYFHWVREBEFVCEPPLIQTHTKLLV 343  
QY 301 LEGQAATLKCAIGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDFITTSQDSGAFTCI 360  
Db 344 LEGQAATLKCAIGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDFITTSQDSGAFTCI 403  
QY 361 AANAAGEATAMVEVSIVQLPHLSNSTRTAPKRSLSIDITGSKTSRGGGGGGGEPPPKS 420  
Db 404 AANAAGEATAMVEVSIVQLPHLSNSTRTAPKRSLSIDITGSKTSRGGGGGGGEPPPKS 463  
QY 421 PPERAVLVSEVTTTSALVKWSKSPRVKMYOLOVNCSDDEVLIYRMIPASNKAFAVNN 480  
Db 464 PPERAVLVSEVTTTSALVKWSKSPRVKMYOLOVNCSDDEVLIYRMIPASNKAFAVNN 523  
QY 481 LVSGTGYDLCLVAMWDDTATTATNIVGCAOFFTKADYPQCSMHSQILGGMILVIGG 540  
Db 524 LVSGTGYDLCLVAMWDDTATTATNIVGCAOFFTKADYPQCSMHSQILGGMILVIGG 583  
QY 541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAASVNVYQTNAGQPPPPSSAPAGAPQGP 600  
Db 584 IIVATLLVFIVILMVRYKVCNHEAPSKMAAASVNVYQTNAGQPPPPSSAPAGAPQGP 643  
QY 601 KVVVRNELDFTASLARASDSSSSSLGSGEAGLGRAPWIPPSAPRPKPSIDRLMGAF 660  
Db 644 KVVVRNELDFTASLARASDSSSSSLGSGEAGLGRAPWIPPSAPRPKPSIDRLMGAF 703  
QY 661 ASLDLSQRKEELLDSTPAGRGAGTSARGHSDREPLGPPAARASLLPLPLEGKAKR 720  
Db 704 ASLDLSQRKEELLDSTPAGRGAGTSARGHSDREPLGPPAARASLLPLPLEGKAKR 763  
QY 721 SHSFDMGDFAAAAGGVPGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARTFGS 780  
Db 764 SHSFDMGDFAAAAGGVPGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARTFGS 823  
QY 781 SEMWMESTV 789  
Db 824 SEMWMESTV 832

RESULT 2

Q9BE71

ID Q9BE71

AG Q9BE71;

PRELIMINARY; PRT; 789 AA.

DT 01-JUN-2001 (TremBirel. 17, Created)  
DT 01-JUN-2001 (TremBirel. 17, Last sequence update)  
DT 01-DEC-2001 (TremBirel. 19, Last annotation update)  
DE Hypothetical 84.7 kDa protein.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=FRONTAL LOBE LEFT;  
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
SU Suzuki Y., Sugano S., Hashimoto K.;  
RT "Isolation of full-length cDNA clones from macaque brain cDNA  
libraries";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB056799; BAB39323.1; -  
DR HSSP; P56276; ILLK.  
DR InterPro; IPR003961; FN III.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR003598; IG\_C2.  
DR InterPro; IPR003006; IG\_MHC.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_Typ.  
DR Pfam; PF00041; fn3; 1.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00560; LRR; 7.  
DR Pfam; PF01463; LRRCT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00060; FN3; 1.  
DR SMART; SM00409; IG; 1.  
DR SMART; SM00408; IG\_C2; 1.  
DR SMART; SM00370; LRR; 5.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 6.  
KW Hypothetical protein; Immunoglobulin domain.  
SQ SEQUENCE 789 AA; 84730 MW; BB86DE81BC284B23 CRC64;

Query Match 99.5%; Score 4075; DB 6; Length 789;

Best Local Similarity 99.4%; Pred. No. 1.5e-308;

Matches 784; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 METLLGGLLAFGMAFAVVDACPKYCVQNLSGLTLCPSKGLLFVPPDIDRRTVELRLG 60  
Db 1 METLLGGLLAFGMAFAVVDACPKYCVQNLSGLTLCPSKGLLFVPPDIDRRTVELRLG 60  
QY 61 GNFIHISRODFANMTGLVDLTLSRNTISHIQPFSLDLESRLSHLDSNRLPSLGEDTL 120  
Db 61 GNFIHISRODFANMTGLVDLTLSRNTISHIQPFSLDLESRLSHLDSNRLPSLGEDTL 120  
QY 121 RGLVNLQHLI VNNNOLGGIADAEFDLTLEDLDSYNNLHGLPWSVRRMVLHQLSL 180  
Db 121 RGLVNLQHLI VNNNOLGGIADAEFDLTLEDLDSYNNLHGLPWSVRRMVLHQLSL 180  
QY 181 DHNLDDHIAEGTFADLQKLARLDLTLSNRLOKLPDPPIFARSOASALTATPPAPPLSFSG 240  
Db 181 DHNLDDHIAEGTFADLQKLARLDLTLSNRLOKLPDPPIFARSOASALTATPPAPPLSFSG 240  
QY 241 GNPLHNCCELLWLRLERDDDLTCTGSPGGLKGRYFHWVREBEFVCEPPLIQTHTKLLV 300  
Db 241 GNPLHNCCELLWLRLERDDDLTCTGSPGGLKGRYFHWVREBEFVCEPPLIQTHTKLLV 300  
QY 301 LEGQAATLKCAIGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDFITTSQDSGAFTCI 360  
Db 301 LEGQAATLKCAIGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDFITTSQDSGAFTCI 360  
QY 361 AANAAGEATAMVEVSIVQLPHLSNSTRTAPKRSLSIDITGSKTSRGGGGGGGEPPPKS 420



```

Db 361 ANAAGEATATVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGSGGGGEPKPS 420
QY 421 PPERAVLVSEVTTTTSALVKWSVSKAPRYMYOLOVNCSDDEVLYRMI PASNKAFVNN 480
Db 421 PPERAVLVSEVTTTTSALAKWSVSKTPRYMYOLOVNCSDDEVLYRMI PASNKAFVNN 480
QY 481 LVSGTGYDLCVLAMWDDTATTTATNIVGCAQFFTKADYPOCOSMHSQILGTMILVIGG 540
Db 481 LVSGTGYDLCVLAMWDDTATTTATNIVGCAQFFTKADYPOCOSMHSQILGTMILVIGG 540
QY 541 IIVATLLVFIIVLMVRYKVCNHEAPSKMAAAVSNVYQTNAGQPPPPSSAPAGAPQGP 600
Db 541 IIVATLLVFIIVLMVRYKVCNHEAPSKMAAAVSNVYQTNAGQPPPPSSAPAGAPQGP 600
QY 601 KVVVRELLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
Db 601 KVVVRELLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
QY 661 ASLDLKSQRKEELDSRTAPAGAGTSGARGHSDREPLIGPPAARARSLLPLPLEGKAKR 720
Db 661 ASLDLKSQRKEELDSRTAPAGAGTSGARGHSDREPLIGPPAARARSLLPLPLEGKAKR 720
QY 721 SHSFDWGFAPAAAAGVPGGYSPPRKVSNITKSLSVNGMLLPFEESDLVGAGTGTGS 780
Db 721 SHSFDWGFAPAAAAGVPGGYSPPRRVSNITKSLSVNGMLLPFEESDLVGAGTGTGS 780
QY 781 SEWNVESTV 789
Db 781 SEWNVESTV 789

RESULT 3
Q9CYK3
ID Q9CYK3 PRELIMINARY; PRT: 788 AA.
AC Q9CYK3;
DC 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE 5730420005RIK protein.
GN 5730420005RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085560; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL; AK017594; BAB30828.1; -.
DR HSSP; P56276; 1TLK.
DR MGD; MGI:1917780; 5730420005RIK.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003599; IG.

```

```

DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00560; LRR; 7.
DR Pfam; PF01463; LRRCT; 1.
DR PRINTS; PR00019; LEURICRPT.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 6.
KW Immunoglobulin domain.
SQ SEQUENCE 788 AA; 85076 MW; 70B70C5E6B52277B CRC64;

Query Match 94.9%; Score 3885.5; DB 11; Length 788;
Best Local Similarity 94.9%; Pred. No. 8.6e-294;
Matches 750; Conservative 10; Mismatches 27; Indels 3; Gaps 2;

QY 1 METLLGGLLAFQMAFVVDACPKYCVQNLSGLTLCPSKGLLPVPPDIDRRTVELRLG 60
Db 1 METLLGGLLAFQMAFVVDACPKYCVQNLSGLTLCPSKGLLPVPPDIDRRTVELRLG 60
QY 61 GNFIHISRODPFANMTGLVDLTLSNTTISHIQPFPSFLDLESURSLHLSNRLPSLGEDTL 120
Db 61 GNFIHISRODPFANMTGLVDLTLSNTTISHIQPFPSFLDLESURSLHLSNRLPSLGEDTL 120
QY 121 RGLVNLQHLVNNNOLGGIADAEFDFLLTLEDLDSYNNHGLPWSVRRMNLHQLSL 180
Db 121 RGLVNLQHLVNNNOLGGIADAEFDFLLTLEDLDSYNNHGLPWSVRRMNLHQLSL 180
QY 181 DNNLDHTAEGTFADLQKLARDLTLSNRLQKLPPDPIFARSOASALTATPPAPPLSFSG 240
Db 181 DNNLDHTAEGTFADLQKLARDLTLSNRLQKLPPDPIFARSOASALTATPPAPPLSFSG 240
QY 241 GNPLHCNCELLWLRELRDDELTCGSPGGLKGRYFVHVRBEFVCEPLITQHTHKLIV 300
Db 241 GNPLHCNCELLWLRELRDDELTCGSPGGLKGRYFVHVRBEFVCEPLITQHTHKLIV 300
QY 301 LEGQAATLKCAIGDPPSLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
Db 301 LEGQAATLKCAIGDPPSLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
QY 361 ANAAGEATATVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGSGGGGEPKPS 420
Db 361 ANAAGEATATVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGSGGGGEPKPS 420
QY 421 PPERAVLVSEVTTTTSALVKWSVSKAPRYMYOLOVNCSDDEVLYRMI PASNKAFVNN 480
Db 421 PPERAVLVSEVTTTTSALVKWSVSKAPRYMYOLOVNCSDDEVLYRMI PASNKAFVNN 480
QY 481 LVSGTGYDLCVLAMWDDTATTTATNIVGCAQFFTKADYPOCOSMHSQILGTMILVIGG 540
Db 481 LVSGTGYDLCVLAMWDDTATTTATNIVGCAQFFTKADYPOCOSMHSQILGTMILVIGG 540
QY 541 IIVATLLVFIIVLMVRYKVCNHEAPSKMAAAVSNVYQTNAGQPPPPSSAPAGAPQGP 599
Db 541 IIVATLLVFIIVLMVRYKVCNHEAPSKMAAAVSNVYQTNAGQPPPPSSAPAGAPQGP 599
QY 600 PKVVVRELLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPRPKPSLDRLMCA 659
Db 600 PKVVVRELLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPRPKPSLDRLMCA 659
QY 660 FASLDLKSQRKEELDSRTAPAGAGTSGARGHSDREPLIGPPAARARSLLPLPLEGKAK 719
Db 660 FASLDLKSQRKEELDSRTAPAGAGTSGARGHSDREPLIGPPAARARSLLPLPLEGKAK 719
QY 720 RSHSFDWGFAPAAAAGVPGGYSPPRKVSNITKSLSVNGMLLPFEESDLVGAGTGTG 779

```

```
Db 721 RHSEDFMGFAAAA--AVPGGYSPPRRVNIWTKRSLVNGMLLPFEESDLVGARGTFG 778
Qy 780 SSEWNESTV 789
Db 779 SSEWNESTV 788

RESULT 4
Q96NI6 PRELIMINARY; PRT; 719 AA.
ID Q96NI6
AC Q96NI6
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE CDNA FLJ30803 fis, clone FEBRA2001245, weakly similar to NAG14.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furiya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamiyama K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Maeuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055365; BAB70910.1;
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00041; fn3.1.
DR Pfam; PF00047; ig.1.
DR Pfam; PF00560; LRR; 6.
DR Pfam; PF01463; LRRCT; 1.
DR PROSITE; PS00037; MYB.1; UNKNOWN.1.
SQ SEQUENCE 719 AA; 79472 MW; CE301F219705FF4B CRC64;

Query Match 44.5%; Score 1821; DB 4; Length 719;
Best Local Similarity 51.4%; Pred. No. 4.6e-133;
Matches 375; Conservative 118; Mismatches 195; Indels 42; Gaps 13;

Qy 1 METLLGLLAFGMAFAVADACPKYCVQNLSESLGTLCPKGLFPVPPIDRRTVELRUG 60
Db 1 MEKILFLYFLGIA-VKAQICPKRCVCQLSPNLATLCAKGLLFPVPPIDRRTVELRLA 59

Qy 61 GNFIHISQDFANMTGLVDLTLSRNTISHIOPFFSLDESLSRLSLDNLPSLGEDTL 120
Db 60 DNFVTNIKKDFANMTSLVDLTLSRNTISFITPHAFADLRNLRALHLSNRLTKITDMF 119

Qy 121 RGLVNLQHLIVNNOLGGIADAFEDFLITLEDLSLYNNHGLPWSVRRMNLHLSL 180
Db 120 SGLSNLHLIINNQLTSLISSTAFDD-VFALBELDLSYNNLETIPMDAVEKWSLHLSL 178

Qy 181 DNNLPHIAEGTFADLQKLARDLTLSRLQKLPDPPIFARSOASALTATPFAPLPSFSG 240
Db 179 DNMINDINPKGTFSHLHKRLDVTNSKLOKLPDPDLFORAQVLATSGIISSTALSFG 238

Qy 241 GNPLHCNCELLMRLRLERDDDLTCGSPGGLKGRYFWHVREBEFVCEPPLITQHTKLIV 300
Db 239 GNPLHCNCELLMRLRLSREDDLETSCASPPLITGRYFWSIPEBEFLCEPPLITRHTHEMRV 298

Qy 301 LEGOATLCKCAIGDPSPIHVVAPDDRLVGNSSRTAVVDNGTLDFITTSQDSGAFTCI 360
Db 299 LEGQATLCKARGDPEPAIHWSISPEGLKISNATSLVYDNGTLDFILITTVKDTGAFTCI 358
```

```
Qy 361 AANAAGEATAMVEVSIVOLPHLSNSTSTAPPKRLSLDITGSSKTSRGGSGGGEPKPS 420
Db 359 ASNPAGEATQIVDLHIKPLHLLNSTNHIHEPDGSSDISTSKSGNSTSSNG---DTK 415

Qy 421 PPERAVLVSEVTTTSALVKWSVKSAPRVKMYOIQYNCSDDDEVLIYRMI PASNKAFVNN 480
Db 416 LSQDKIVVAEATSTALLKFNQIRNPGIRFQIYNGTYDDTLVYRMIPTSKTFLVNN 475

Qy 481 LVSGTGYDLCLVAMWDDTATLTATNIVGCAQFFTKADYPOCOSMHSOILGGMILVIGG 540
Db 476 LAAGTMYDLCLVALYDDGITSLTATRVVGCILQFTTEQDYVRCHFMOSFGLGTMIIIGG 535

Qy 541 IIVATLLVFIILVMRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPLPPSPAPAGPQGP 600
Db 536 IIVASVLVFIILMIRYKVCNNNGQHKV-TKVSNNVYSQTNGAQ-----IQGCVTLTPQSVS 590

Qy 601 KVVVRNELLDPTASLARASDSS---SSSSLSGGEAAGLRAPWRIPPSAPRPKPSLDRLM 657
Db 591 KQAVGHE---ENAOCCATSDNVIOSSSETCSSQSSTTTSA---LPPS-----W 633

Qy 658 GAFASLDLKSQRKELLDSRTPAGRGAGTSARGHSDREPLLGPPAARASLLPLP---- 713
Db 634 TSSTSVSQKQRKTGKTKEPQNE-AVTNVESQNTNRN-----NSTALQLVSRPPDSV 686

Qy 714 LEG-KAKESH 722
Db 687 TEGPSTSKRAH 696

RESULT 5
Q9BTNO PRELIMINARY; PRT; 628 AA.
ID Q9BTNO
AC Q9BTNO
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-DEC-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical 66.3 kDa protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003578; AAH03578.1;
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR00483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00041; fn3.1.
DR Pfam; PF00047; ig.1.
DR Pfam; PF00560; LRR; 6.
DR Pfam; PF01463; LRRCT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00060; FN3.1.
DR SMART; SM00409; IG.1.
DR SMART; SM00408; IGC2.1.
DR SMART; SM00082; LRRCT.1.
DR SMART; SM00013; LRRNT.1.
DR SMART; SM00369; LRR_TYP.6.
KW Hypothetical protein; Immunoglobulin domain.
SQ SEQUENCE 628 AA; 66259 MW; FEDC5A5056ABC5FC CRC64;

Query Match 38.4%; Score 1573; DB 4; Length 628;
Best Local Similarity 53.1%; Pred. No. 7.8e-114;
Matches 312; Conservative 88; Mismatches 174; Indels 14; Gaps 5;
```

QY 16 AVADACPKYCVQCNLSESLGTLCPKSKGLLFVPPDIDRRTVELRLGGNFIIHISRODFANM 75  
 Db 23 ATPSPCPRCRCQOTOSLPLSLCPGAGLLFVPPSLDRRAAEELRADNEFIASVRRDLANM 82  
 QY 76 TGLVDLTNRNTISHIQPFSLDLLESRLSLHDSNRLSLGDETLRLGLVNLQHLIVNNQ 135  
 Db 83 TGLHLLSRNTIRHVAAGAFADRALRALHLDGNRLTSLGEGQLRGVLNRLHLILSNQ 142  
 QY 136 LGGIADEAFEDFLITLEDLDSYNHLGLPWSVRRMVLNHLQSLSDHNLHLHIAEGTFAD 195  
 Db 143 LAALAAGALDDCAETLEDLDSYNHLGLPWEALGRLGNVNTLGLDHNLLASVPAGAFSR 202  
 QY 196 LOKLARLDLTNRLOKLPDPIFARSOASATATPFPAPLS---PSFGGNPLHNCNELLW 252  
 Db 203 LHLKARLDMTNRLLTTPDPLFSR--LPLARPRGSPASALVLAFCGNPLHNCNELLW 259  
 QY 253 LRLERDDLTETCGSPGGLKGRYFWHVREEFVCEPPLITQHTHKLVLLEGOAATLKCA 312  
 Db 260 LRLAREDDLEACASPPALGGRYFWAGVEEFVCEPVPVTHRSPPPLAYPAGRPALRCRA 319  
 QY 313 IGDSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQSGAFTCIAANAAGEATAMV 372  
 Db 320 VGDPEPRVWSPQGRLLGNSSRARAFNGTLELLVTEPGDGIPTCIAANAAGEATAV 379  
 QY 373 EVSI--VOLPHLSNSTSRTAPPKSRSLDITGSSKTSRGGGGGGEPPKSPPERAVLYSE 430  
 Db 380 ELTVGPPPPPOLANSTCDPPRGDGDALTTPPSAASASAKVADTG---PPTDRGVQVTE 435  
 QY 431 VTTTSALVKNWSKSAKAPVKMYQOYNCSDDDEVLYRMIIPASNAKAFVNNVLVSGTYDLC 490  
 Db 436 HGATAALVQWPDQRPPIGRMYQIQYNSADDILYRMIIPAESRFLTLASGRTYDLC 495  
 QY 491 VLAWDDTATLTATNIVGCAOFFTKADYPOCOSMHSOILGTMILVIGGIIVATLLVFI 550  
 Db 496 VLAVYEDATGLTATRPVGCARFSTEPALRPGCAHPAFELGTMILVIGGIIVASVLVFI 555  
 QY 551 VILVMRYKVCNHEAP--SKMAAAVSNVYQTNAGQPPPPSSAPGAPP 596  
 Db 556 FVLLMRYKVGQPPGKAKIPAPVSSVCSQTNALGTPPTPAPPAPPEP 603

RESULT 6  
 Q9P244 ID Q9P244 PRELIMINARY; PRT; 700 AA.  
 AC Q9P244;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE K1AA1484 protein (Fragment).  
 GN K1AA1484.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20277482; PubMed=10819331;  
 RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human  
 genes.XVII.The complete sequences of 100 new cDNA clones from brain  
 RT which code for large proteins in vitro.";  
 RL DNA Res. 7:143-150(2000).  
 DR EMBL: AB040917; BAA96008.1; -;  
 DR InterPro: IPR003961; FN\_III.  
 DR InterPro: IPR003598; IG\_C2.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR000483; LRR\_Cterm.  
 DR InterPro: IPR003592; LRR\_Out.  
 DR InterPro: IPR003591; LRR\_Typ.  
 DR InterPro: IPR001005; Myb\_DNA\_binding.  
 DR Pfam: PF00041; fn3; 1.

DR Pfam: PF00047; ig; 1.  
 DR Pfam: PF00560; LRR; 6.  
 DR Pfam: PF01463; LRRCT; 1.  
 DR SMART: SM00060; FN3; 1.  
 DR SMART: SM00408; IG2; 1.  
 DR SMART: SM00370; LRR; 4.  
 DR SMART: SM00082; LRRCT; 1.  
 DR SMART: SM00369; LRR\_Typ; 1.  
 DR PROSITE: PS00037; MYB\_1; UNKNOWN\_1.  
 KW Immunoglobulin domain.  
 FT NON\_TER  
 SQ SEQUENCE 700 AA; 74865 MW; 30D0432242494159 CRC64;  
 Query Match 38.4%; Score 1571.5; DB 4; Length 700;  
 Best Local Similarity 45.7%; Pred. No. 1.2e-113;  
 Matches 352; Conservative 106; Mismatches 203; Indels 109; Gaps 18;  
 QY 59 LGGNFIIHISRODFANMTGLVDLTLSRNTISHIQPFSLDLLESRLSLHDSNRLSLGDE 118  
 Db 1 LTDNFIAAVRRDRFANMTSLVHLTSLRNTICQVAGAFADRALRALHLDNRRLAEVRGD 60  
 QY 119 TLRGLVNLQHLIVNNNLGGIADEAFEDFLITLEDLDSYNHLGLPWSVRRMVLNHLQ 178  
 Db 61 QLRLGLNLRHLILGNNOIRRVESAAFDALSTVEDLDSYNHLGLPWEALGRLGNVNTL 120  
 QY 179 SLDNHLHDHIAEGTFADLQKLARLDLTNRLOKLPDPIFARSOASATATPFPAP-PLSF 237  
 Db 121 TLDNHLHDHIAEGTFVQLHKLVRDMLTNRLOKLPDGLFLRSQGTG---PKPTPLTV 176  
 QY 238 SFGGNPLHNCNELLWRLERDDLTETCGSPGGLKGRYFWHVREEFVCEPPLIT-QHTH 296  
 Db 177 SFGGNPLHNCNELLWRLERDDLTETCATDEHLDTRYFWSIPEEFVCEPPLITRQAGG 236  
 QY 297 KLLVLEGOAATLKCAIGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGA 356  
 Db 237 RALVVEGOAVSLRCRAVGDPEPVHVAWAPDORLLGNSSRTVRGDTLDTITTLRDSGT 296  
 QY 357 FTCTAANAAGEATAMVEVSIIVOLPHLSNSTSRTAPPKSRSLDITGSSKTSRGGGGGGE 416  
 Db 297 FTCTAANAAGEATAPVEVCVVPLPLM-----APPPAAPPLTEPGSSDIATPCRPGAN 349  
 QY 417 PKSPPERAVLYSEVTTTSALVKNWSKSAKAPVKMYQOYNCSDDDEVLYRMIIPASNAKAF 476  
 Db 350 --DSAAERLVAELTNSVLIRWPAQRPVGRMYQYNSVDDSLVYRMIPTSTQTF 407  
 QY 477 VNNLVSGTYDLCVLAWDDTATLTATNIVGCAOFFTKADYPOCOSMHSOILGCTMIL 536  
 Db 408 LVNDLAGRAYDLCVLAVYDDGATALPATRVVGVQVFTTAGDPAPCRPLRAHFLGCTMII 467  
 QY 537 VIGGIIVATLLVFIIVLMVRYKV-----CNHEAPSKMAAAVSNVYQTNAGQPPPPSSAP 591  
 Db 468 AIGGVIVASVLVFIIVLMIRYKVYGDGDSRRVKGRSLPRVSHVCSQTNGA--GTGAQ 524  
 QY 592 AGAPP-----QGPVKVVRNELLDTFTASLARADSSSSSSSLGSGEAAAGLGRA 638  
 Db 525 APALPAQDHYEALREVESQAAPAVAVE-----AKAMEAETASA----- 562  
 QY 639 FWRIPPSAPRPKPSLDRLMGAFASLDLKSQRKEELDDSRTPAGRGACTSARGHSDREP- 697  
 Db 563 -----EPEVILGRSLGGSATSLCLLPSSEETSGEESR-----AAGVPRRSRG 604  
 QY 698 LGLGPPAARARSLPLP-----LEGKAKRSHSPDMGDFAAAAAGGVVPGVYPPRKYSNLWT 753  
 Db 605 ALEPPTSAPPTALVPGGAAARPPQORYSFD-GDY-----GALFQSHSYPRRARR--T 655  
 QY 754 KRSLSVNGMLLPFEESDLVGARG-----TFGSEWVWMESTV 789  
 Db 656 KHRST-----PHLDGAGGAAGEDGDLGLGSARACLAFTSTEWMLLESTV 700

RESULT 7  
 Q99KT6 ID Q99KT6 PRELIMINARY; PRT; 492 AA.



Q96CS0		PRELIMINARY;	PRT;	450 AA.
ID	Q96CS0			
AC	Q96CS0;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)			
DE	Hypothetical 47.4 kDa protein (Fragment).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
ON	NCBI_TaxID=9606;			
OX	[1]			
RP	SEQUENCE FROM N. A.			
RC	TISSUE=OVARY;			
RA	Strausberg R.;			
RL	Submitted (Oct-2001) to the EMBL/GenBank/DBDJ databases.			
DR	EMBL; BC014678; AAH14678.1; ..			
DR	InterPro; IPR001961; FN III.			
DR	InterPro; IPR001006; Ig_MHC.			
DR	Pfam; PF00041; fn3; 1.			
DR	Pfam; PF00047; ig; 1.			
DR	Hypothetical protein.			
FT	NON TER			
FT	NON TER			
SEQ	SEQUENCE 450 AA; 47417 MW; 93CEE7B44D2C1F86 CRC64;			
Query Match 18.9%; Score 775; DB 4; Length 450;				
Best Local Similarity 38.8%; Pred. No. 6.7e-52;				
Matches 189; Conservative 75; Mismatches 138; Indels 84; Gaps 13;				
QY	297	KLIVLEQQAATLKCAIKGDSPLIHVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGA	356	
DB	6	RALVVEQQAASLCRAVYGDPEVPHVHVPDGRLLGNSRTRVRGDTGLDVTITLRDST	65	
QY	357	FTCIAANAAGEATAMVSVISVOLPHLSNSTSRTPAPKSRSLSDITGSSKTRGGGSGGGE	416	
DB	66	FTCIASNAAGEATAPVEVCVPLPM-----APPPAAPPLPTEPGSSDIATPERPCAN	118	
QY	417	PPKSPPERAVLVSEVTTTTSALVKWSKSAAPRVRYQLQVNCSDDEVLIVRMIPASNKAF	476	
DB	119	--DAAERRLVAELTNSVLIRWPAQRPVPGIRMYQVQVNSSVYDLSLVYRMIFSTQTF	176	
QY	477	VVNLVSGTYDLCVLAWDDTATTLATNIVGCAQFTKADYQCQSMHSQILGGTMIL	536	
DB	177	LVNDLAAAGRAYDLCVLAVYDDGATALPATRVYGCVFQTTAGDPAPCRPLRAHFLGGTMII	236	
QY	537	VIGIIVATLLVFIVILMVRYKV-----CNHEAPSKMAAASVNVYSTNGAQPPPSSAP	591	
DB	237	AIGGVIVASVLVFLVLLMIRYKVYDGDGSRVRKGSRLPRVSHVCSTNGA---GTGAQAQ	293	
QY	592	AGAPP-----QQPPKVVVRNELLDTTASLARASDSSSSSLGSGEAAAGLGRA	638	
DB	294	APALPAQDHVEALREVESQAAPAVE-----AKAMEAETASA-----	331	
QY	639	PWRIPPSAPRPKSLDRLMGAFA-SLDLKSQRKEELDSRTPAGRGAGTSARGHHSREP	697	
DB	332	-----BPEVVLGRSLGSGATSLSCLLPSEETSGBESR-----AAVGPRRRSRG	373	
QY	698	LLGPPAARASLLPLP-----LEGKAKSHSDFMDGFAAAAGGVVPGYSPPRKVSNIWT	753	
DB	374	ALPEPTSAPPTLALVPGAAARPPQORYSFD-GDY-----GALFOSHYSYPRRARR--T	424	
QY	754	KRSLs 758		
DB	425	KRHS 429		
RESULT 10				
Q9BWJ0	Q9BWJ0	PRELIMINARY;	PRT;	324 AA.
ID	Q9BWJ0			
AC	Q9BWJ0;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)			

Hypothetical 33.2 kDa protein.  
 DE Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=EYE;  
 RA Strausberg R.;  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=COLON;  
 RA Strausberg R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LUNG;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC000207; AAH0207.1; -  
 DR EMBL; BC014040; AAH14040.1; -  
 DR EMBL; BC015581; AAH15581.1; -  
 DR InterPro; IPR003961; FN\_III.  
 DR Pfam; PF00041; fn3; 1.  
 DR SMART; SM00060; FN3; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 324 AA; 33192 MW; C84CCC9A811E877D CRC64;

Query Match 16.0%; Score 653; DB 4; Length 324;  
 Best Local Similarity 34.1%; Pred. No. 1.3e-42;  
 Matches 161; Conservative 54; Mismatches 107; Indels 150; Gaps 10;

Qy 320 IHWVAPDDRLLVGNSSRTAVYDNGTLDFITTSODSGAFTCIAANAAGEATAMVEVSIVQL 379  
 Db 1 MHWVGPDRLVGNSSRARAFNGTLGIVTGADAGGYCIATNPAGEATARVELRVLAL 60

Qy 380 PHLSNSTSRAPPKSLSDITGSSKTSRGGGSGGGEPPKSPPERAVLVSEVTTTSALVK 439  
 Db 61 PHGGNSSAEGGRPGP--SDIAASARTAAEGGTLESEP-----AVQVTEVTATSGLVS 111

Qy 440 WSVKSAAPRVKMYQLQYNGSDDEVLYRIMPASNKAFFVNVLVSGTGYDLCLVAMWDDTA 499  
 Db 112 WGFGRPADPVMMFQIQYNSSDEDTLIYRVPASSHHFLKHLVPGADYDLCLLALSPAAG 171

Qy 500 -TTLTATNIVGCAQFTKADYPCQSMHSQILGTMILVIGGIIVATLFIIVILMVRYK 558  
 Db 172 PSDLTATRLGCAHFSTLPASPCHALQAHVLGTLTVAVGVLVAALLVFTVALLVRGR 231

Qy 559 -VCNHEAPSKMAAASNVYSQTNGAOPPPSSAPAGAPQGPVKVYVRNELLDFTASLAR 617  
 Db 232 GAGNGRLPKL-----SHVQSQTNGG---PSPTPKAHPRSP----- 266

Qy 618 ASDSSSSSLGSGEAGLGRAPWRIPSPAPRPKPSLDRLMGAFSLDLKSQRKEELDSR 677  
 Db 267 -----PRP----- 269

Qy 678 TPAGRGAGTSARGHSHSDREPLLCPPAARASLLPLPLEGKAKESHSDMGDFAAAGGV 737  
 Db 270 -----QKSCSLDLDG----- 279

Qy 738 VPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTGFSGSEWMESTV 789  
 Db 280 -ACYGYARRLGGAWARRSHSVHGGLLG-----ACRGVGGSAERLEESVV 324

RESULT 11  
 Q9HCJ2 PRELIMINARY; PRT; 640 AA.  
 ID Q9HCJ2  
 AC Q9HCJ2; 16, Created)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)

DE	KIAA1580 protein (Fragment).	QY	363	NAAGATAMVEVSIVQLPHLSNSTSRTAPPKRSKLSIDITGSKTSRGGGSGGEPKSP	422
GN	KIAA1580	Db	430	NSVGNVTASATL-----NVTAATTPFSFSTV	458
OS	Homo sapiens (Human)	QY	423	ERAVLVSEVTTTSALVKWSKSAAPRVKMYQLQYNCSDDEVLIYRMI PASNAKAFVNNLV	482
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Db	459	-----VETMEFSQDEARTDNNV	476
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	QY	483	SGTGYDLCVLAMWDDT--ATTLTATNIVGCAQFFT-----KADYPQCSMHQSILGGTM	534
NCBI_TaxID=9606;		Db	477	GPT-----PVDWETNTVTSITPQSTRSTKTFITPVTIDNSGIPGI-----DEVNKTITK	527
[1]		QY	535	ILVIGGIIVATLLVFIVILMVRYK-----CNHAPSKMAAAVSNVVSQTNGAOP----	584
RP	SEQUENCE FROM N.A.	Db	528	III--GCFVAITLMAAAMLVIFVKRKQHHQHPTR--TVEIINVDDEITGDTPMESH	584
RX	MEDLINE=20450683; PubMed=1097877;	QY	585	-----PPPS 588	
RA	Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.,	Db	585	LPMPA 589	
RT	"Prediction of the coding sequences of unidentified human genes.				
RT	xviii. The complete sequences of 100 new cDNA clones from brain which				
RT	code for large proteins in vitro."				
RL	DNA Res. 7:273-281(2000).				
DR	EMBL; AB046800; BAB13406.1; --				
DR	HSSP; P22888; ILUT				
DR	InterPro; IPR003599; IG.				
DR	InterPro; IPR003598; IG_c2.				
DR	InterPro; IPR003600; IG_like.				
DR	InterPro; IPR003006; IG_MHC.				
DR	InterPro; IPR001611; LRR.				
DR	InterPro; IPR000483; LRR_Cterm.				
DR	InterPro; IPR000372; LRR_Nterm.				
DR	InterPro; IPR003592; LRR_Out.				
DR	InterPro; IPR003591; LRR_Typ.				
DR	Pfam; PF00047; IG; 1.				
DR	Pfam; PF00560; LRRCT; 1.				
DR	Pfam; PF01463; LRRCT; 1.				
DR	Pfam; PF01462; LRRNT; 1.				
DR	PRINTS; PR00019; LEURICHRPT.				
DR	SMART; SM00409; IG; 1.				
DR	SMART; SM00408; IGC2; 1.				
DR	SMART; SM00410; IG_like; 1.				
DR	SMART; SM00370; LRR; 6.				
DR	SMART; SM00082; LRRCT; 1.				
DR	SMART; SM00013; LRRNT; 1.				
DR	SMART; SM00369; LRR_Typ; 8.				
KW	Immunoglobulin domain.				
FT	NON TER				
SQ	SEQUENCE 640 AA; 71949 MW; 6058974872636838 CRC64;				
Query Match					
Best Local Similarity 10.6%; Score 435.5; DB 4; Length 640;					
Matches 163; Conservative 87; Mismatches 238; Indels 177; Gaps 19;					
QY	2	BTLLGLLAFQMAFAV-----VDACPKYVCVCONLSESLGTLCPKGLLFVPPDIDRRIVEL	57		
Db	24	DPLLWLLALQLLVVAGLVRAQTCPSCVCSNQFSKV--ICVRKNLREVDPDGISTNTRL	81		
QY	58	RLGNGFIHISRQDPANMTGLVLTLSRNTISHIQPFSLDLSRLSHLDSNRLPSLGE	117		
Db	82	NLHENQIOLIKVNSFKHURHLEILQLSNRHIRTIEIGAFNGLANLNTLELFDNRLTTPN	141		
QY	118	DTLRGLVNLQHLVNNN-----QLGGTADAEAFEDF-----	147		
Db	142	GAFVYLSKLKELWLRNNPIESIPYAFNRIPSLRRLDGLKRLSYISEGAFGLSNLRY	201		
QY	148	-----LLTLEDLDSYNNLHGLPWSVRRMVNHLQSLDHLNLDHIAEG	191		
Db	202	LNLAMCNLREIPNLTPFLKLDLSDLSGNHLSAIRPGSQGLMHKLQKLMQISOQVIERN	261		
QY	192	TFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLISFSGGNPLHCNCELL	251		
Db	262	AFDNLQSLVEINLHNNLTLLPHD--LF-----TPLHHLERILHLHNPWNCNDIL	310		
QY	252	WLRLERD-----DDLETCSGPGGLKGRYFVHVEEBEVECPPLITQHTKLLVLEGOA	305		
Db	311	WLSWTKDMAPNSNTACCAFCNTPPNLKGRYIGELDQNTFTCVAPVIVEPPADLVNTEGMA	370		
QY	306	ATLKCAKAGDPSPLIHVAPDRDLVGNSS---RTAVYDNGTLIDFITTQSQSGAFTCAA	362		
Db	371	AEKCRG--STSLTSVSWITPENGTVMTGAYKVRIVLSDGTLNFTNTVTDGTGMYTCMS	429		

RESULT 12

Q9BZ20

ID Q9BZ20 PRELIMINARY; PRT; 606 AA.

AC Q9BZ20; (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE BA438B23.1 (Neuronal leucine-rich repeat protein) (CDNA FLJ31810 fis,

clone NT2R12009289, weakly similar to carboxypeptidase N 83 kDa

chain).

GN BA438B23.1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Babbage A.;

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Ishibaishi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,

Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,

Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,

Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,

Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,

Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,

Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,

Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;

RA "NEDO human cDNA sequencing project";

RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL; AL353746; CAC22713.1; --

DR EMBL; AK056372; BAB71167.1; --

DR InterPro; IPR003599; IG.

DR InterPro; IPR003598; IG\_c2.

DR InterPro; IPR003600; IG\_like.

DR InterPro; IPR003006; IG\_MHC.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR000372; LRR\_Nterm.

DR InterPro; IPR003591; LRR\_Typ.

DR Pfam; PF00047; IG; 1.

DR Pfam; PF00560; LRR; 10.

DR Pfam; PF01463; LRRCT; 1.

DR Pfam; PF01462; LRRNT; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00409; IG; 1.

DR SMART; SM00408; IGC2; 1.

DR SMART; SM00410; IG\_like; 1.

DR SMART; SM00082; LRRCT; 1.

DR SMART; SM00013; LRRNT; 1.

DR SMART; SM00369; LRR\_Typ; 10.

KW Immunoglobulin domain.  
SQ SEQUENCE 606 AA; 68065 MW; CB608E281B066B9D CRC64;  
Query Match 10.2%; Score 418.5; DB 4; Length 606;  
Best Local Similarity 26.0%; Pred. No. 6.1e-24;  
Matches 125; Conservative 67; Mismatches 157; Indels 131; Gaps 12;

QY 21 CPKYCVQNLSSSLGTLCPKSGLLFPVPPIDRRRTVELRLGGNFIIHISRODFANMTGLVD 80  
DB 46 CPVSCSCSNQFSKV--VCTRRGLSEVPQIGPNSRYLNLMMENNIQMIQADTFRHLHLLEV 103  
QY 81 LTLRNRTISHIQPFSLDLESRLSLHSDNRLPSIGEDTLRGLVNLQHLIIVNNQLGGIA 140  
DB 104 LQIGRNSIRQIEVGAFAFGLASLNTLDFNLWLTIVPSGAFEYLSKLRLWLRNRPESIP 163  
QY 141 DEAFEDF-----DFLLTLEDLDSLYNNHGLPWDSVRRMVNLH 176  
DB 164 SYAFNRVPSLRLDLGELKKLEYISEGAFEGFLNKLNLGMCNKKDMP--NLTPLVGLUE 221  
QY 177 QLSLDHNLDDHTAEGTFADLQKLAR-----LDLTSNRLQKLPPDP 216  
DB 222 ELEMNGHHPPEIRPGSFHGLSSLLKLLWVMSHERNAFDGLASLVELNLAHNNLSSLPHD- 280  
QY 217 IFARQASALTATPFPAPPLSFSGGNPLHCNCELLWLRRLERD---DDLETG---SPGG 270  
DB 281 LF-----TPLYLYVELHLHHNPMNCDCDILWAWLREYIPTNSTCCGRCHAPMH 330  
QY 271 LKGRYFWHREBEFVCEPPLITQHTKLLVLEGOAATLKCAIGDPSLIHWVAPDDRLV 330  
DB 331 MRGRYLVDEVDQAFQCSAFPINDAPRDNLISGRMAELKCR--PPMSSVKWLLPNGTVL 388  
QY 331 GNSG---RTAVYDNGTLDIFITTSQDSGAFTCIAANAAG--EATAMVEVSIIVOLPHLSNS 385  
DB 389 SHASRHPRIISVLDGTLNFSHLLSDTGVYTCWTVNAGNSASAVLNYSTAEL-NTSNV 447  
QY 386 TSSTAPPKRLSDITGSSKTSRGGSGGGGPPKSPERAVLSEVTTTSALVKVSVSKS 445  
DB 448 SFPT-----TWTVTETETIS----- 461  
QY 446 APRVKMYQLQYNCSDDEVLIYRMI PASNKAFFVNNLVSGTGYDCLVLAEMDDTATLTAT 505  
DB 462 -----PDTTRKYKVPPTIS-----TGY-----QPAYTTSTT 488  
QY 506 NIVGCAQFFTKADYP-----CQCSMHSQLGQTMILVIGGIIVATLLVFILVMYRKV 559  
DB 489 VLIQTTRVPKQVAVPATDITDKMQTSLDEVMKTKIII--GCFVAVTLLAAAMLIVFYKL 546  
QY 560 -CNHEAPSMAAAVSNVYQTNGCAQPPPPSSAPAGAPPOG 598  
DB 547 RKRHQQRSTVTAARTVEIIQVD-EDIPAATAAATAAPSG 585

RESULT 14  
Q8V135 PRELIMINARY; PRT; 648 AA.  
ID Q8V135  
AC Q8V135; 2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Brain tumor-associated protein MBAG1.  
GN Mus musculus (Mouse)  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/J;  
RA Wang J., Bin L., Li G.;  
RT "Brain-specific gene, downregulated in brain tumor";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF300458; AAL6761.1;  
DR InterPro; IPR003598; Ig.  
DR InterPro; IPR003598; Ig.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR typ.

KW Immunoglobulin domain.  
SQ SEQUENCE 606 AA; 68065 MW; CB608E281B066B9D CRC64;  
Query Match 10.2%; Score 418.5; DB 4; Length 606;  
Best Local Similarity 26.0%; Pred. No. 6.1e-24;  
Matches 125; Conservative 67; Mismatches 157; Indels 131; Gaps 12;

QY 21 CPKYCVQNLSSSLGTLCPKSGLLFPVPPIDRRRTVELRLGGNFIIHISRODFANMTGLVD 80  
DB 46 CPVSCSCSNQFSKV--VCTRRGLSEVPQIGPNSRYLNLMMENNIQMIQADTFRHLHLLEV 103  
QY 81 LTLRNRTISHIQPFSLDLESRLSLHSDNRLPSIGEDTLRGLVNLQHLIIVNNQLGGIA 140  
DB 104 LQIGRNSIRQIEVGAFAFGLASLNTLDFNLWLTIVPSGAFEYLSKLRLWLRNRPESIP 163  
QY 141 DEAFEDF-----DFLLTLEDLDSLYNNHGLPWDSVRRMVNLH 176  
DB 164 SYAFNRVPSLRLDLGELKKLEYISEGAFEGFLNKLNLGMCNKKDMP--NLTPLVGLUE 221  
QY 177 QLSLDHNLDDHTAEGTFADLQKLAR-----LDLTSNRLQKLPPDP 216  
DB 222 ELEMNGHHPPEIRPGSFHGLSSLLKLLWVMSHERNAFDGLASLVELNLAHNNLSSLPHD- 280  
QY 217 IFARQASALTATPFPAPPLSFSGGNPLHCNCELLWLRRLERD---DDLETG---SPGG 270  
DB 281 LF-----TPLYLYVELHLHHNPMNCDCDILWAWLREYIPTNSTCCGRCHAPMH 330  
QY 271 LKGRYFWHREBEFVCEPPLITQHTKLLVLEGOAATLKCAIGDPSLIHWVAPDDRLV 330  
DB 331 MRGRYLVDEVDQAFQCSAFPINDAPRDNLISGRMAELKCR--PPMSSVKWLLPNGTVL 388  
QY 331 GNSG---RTAVYDNGTLDIFITTSQDSGAFTCIAANAAG--EATAMVEVSIIVOLPHLSNS 385  
DB 389 SHASRHPRIISVLDGTLNFSHLLSDTGVYTCWTVNAGNSASAVLNYSTAEL-NTSNV 447  
QY 386 TSSTAPPKRLSDITGSSKTSRGGSGGGGPPKSPERAVLSEVTTTSALVKVSVSKS 445  
DB 448 SFPT-----TWTVTETETIS----- 461  
QY 446 APRVKMYQLQYNCSDDEVLIYRMI PASNKAFFVNNLVSGTGYDCLVLAEMDDTATLTAT 505  
DB 462 -----PDTTRKYKVPPTIS-----TGY-----QPAYTTSTT 488  
QY 506 NIVGCAQFFTKADYP-----CQCSMHSQLGQTMILVIGGIIVATLLVFILVMYRKV 559  
DB 489 VLIQTTRVPKQVAVPATDITDKMQTSLDEVMKTKIII--GCFVAVTLLAAAMLIVFYKL 546  
QY 560 -CNHEAPSMAAAVSNVYQTNGCAQPPPPSSAPAGAPPOG 598  
DB 547 RKRHQQRSTVTAARTVEIIQVD-EDIPAATAAATAAPSG 585

RESULT 13  
Q9A85 PRELIMINARY; PRT; 649 AA.  
ID Q9A85  
AC Q9A85;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical 72.3 kDa protein.  
GN BAG.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang J.;  
RL Thesis (2000); Zhongshan Medical University, Guangzhou, China.  
DR EMBL; AJ297858; CAC82651.1;  
DR InterPro; IPR003006; Ig MHC.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00560; LRR; 8.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01462; LRRNT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR Hypothetical protein; Signal.  
FT SIGNAL 1 38 POTENTIAL.  
SQ SEQUENCE 649 AA; 72313 MW; 9C11C5ABC7E536CF CRC64;  
Query Match 9.4%; Score 385; DB 4; Length 649;  
Best Local Similarity 23.8%; Pred. No. 2.7e-21;





Db 489 TSTTVLIOTTRVPKQVAVPATDTTDMQTSLEVMKTKIII--GCFVAVTLLAAAMLIV 546  
Qy 556 RYKV-CNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPQG 598  
Db 547 FYKLKRKHQORSTVTAARTVEIIQVD-EDIPAATSAATAAPSG 589

Search completed: June 8, 2003, 21:30:25  
Job time : 69 secs

---

**THIS PAGE BLANK (USPTO)**